

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Estatística

Introdução ao Controle Estatístico de Processo
on-line

Paulo Henrique Ferreira da Silva
Orientador: Prof. Dr. Francisco Louzada Neto

Dissertação apresentada ao Departamento de Estatística da Universidade Federal de São Carlos – DEs/UFSCar, como parte dos requisitos para obtenção do título de Mestre em Estatística.

São Carlos
Março de 2011

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

S586ic

Silva, Paulo Henrique Ferreira da.

Introdução ao controle estatístico de processo *on-line* /
Paulo Henrique Ferreira da Silva. -- São Carlos : UFSCar,
2012.

193 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2011.

1. Estatística. 2. Controle Estatístico do Processo (CEP).
3. Controle de qualidade. 4. Controle de qualidade - gráficos,
tabelas, etc. 5. Capabilidade. I. Título.

CDD: 519.5 (20^a)

Paulo Henrique Ferreira da Silva

**INTRODUÇÃO AO CONTROLE ESTATÍSTICO DE PROCESSO ON-
LINE**

Dissertação apresentada à Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Estatística.

Aprovada em 28 de fevereiro de 2011.

BANCA EXAMINADORA

Presidente _____
Prof. Dr. Francisco Louzada Neto (DEs-UFSCar/Orientador)

1º Examinador _____
Prof. Dr. Benedito Galvão Benze (DEs-UFSCar)

2º Examinador _____
Profa. Dra. Gleici da Silva Castro Perdoná (FMRP-USP)

Dedico esta dissertação, com muito amor e carinho, aos meus pais Paulo Sérgio Ferreira da Silva e Alzira Venturin Ferreira da Silva, que sempre me incentivaram e me ajudaram a lutar pelas coisas em que eu acredito, e ao meu tio Itamar dos Santos Cunha, que sempre acreditou que eu fosse capaz.

"Ah, but a man's reach should exceed his grasp, or what's a heaven for?"

(Robert Browning - Men and Women and Other Poems)

Agradecimentos

Agradeço...

Primeiramente a Deus, que sempre me iluminou e me guiou para mais esta realização em minha vida.

Aos meus familiares, Paulo Sérgio, Alzira, Gustavo, Márcio e Itamar, que sempre estiveram do meu lado.

Ao meu orientador Prof. Dr. Francisco Louzada Neto, por ter acreditado em minha capacidade e ter me dado a oportunidade de iniciar este trabalho.

Ao Prof. Dr. Benedito Galvão Benze e ao Prof. Dr. Francisco Antonio Rojas Rojas, pelas sugestões e comentários dados quando da realização do meu exame de qualificação.

A todos os professores e funcionários do Departamento de Estatística da UFSCar - DEs - pelo apoio às minhas dificuldades e pelo companheirismo.

A todos os amigos que, de alguma maneira, contribuíram para a realização deste trabalho. Em especial, aos amigos, Anderson e Caroline, pela amizade e carinho e pela ajuda no decorrer do trabalho.

Resumo

Neste trabalho são apresentadas algumas ferramentas do Controle Estatístico de Processos (CEP), que podem ser usadas no monitoramento de sistemas produtivos ao longo do tempo, bem como a sua aplicação em conjuntos de dados artificiais, que fazem parte de contextos reais. O estudo sobre essas ferramentas estatísticas está distribuído de tal maneira que são abordados durante o trabalho: a teoria das ferramentas estudadas, os diferentes contextos em que podem ser aplicadas e a sua implementação *on-line*, utilizando recursos computacionais de *softwares* livres. A aplicação no sistema *on-line* é realizada de tal modo que viabiliza a praticidade e eficácia na geração de gráficos para o CEP e de índices que refletem a capacidade do processo. É apresentada também uma sequência de passos para o uso do sistema de CEP *on-line* aqui proposto, tanto para situações em que se observa apenas uma medida da qualidade do processo (caso univariado), quanto para situações em que são observadas várias medidas (caso multivariado).

Palavras-chave: Qualidade, Controle Estatístico de Processos (CEP), Ferramentas CEP, Política de Controle 6-sigma, Gráficos de Controle, Capacidade do Processo, Sistema *on-line*.

Sumário

1 Introdução	5
1.1 Motivação e objetivo.....	6
1.2 Os dados.....	4
1.3 A política de controle 6-sigma.....	5
1.4 Organização do trabalho.....	6
2 Metodologia CEP para variáveis	8
2.1 Gráfico S	11
2.1.1 Gráfico S para amostras de mesmo tamanho.....	11
2.1.1.1 Gráfico S com especificação (amostras de mesmo tamanho).....	12
2.1.1.2 Gráfico S sem especificação (amostras de mesmo tamanho).....	13
2.1.2 Gráfico S para amostras de tamanhos variáveis.....	14
2.1.2.1 Gráfico S com especificação (amostras de tamanhos variáveis).....	14
2.1.2.2 Gráfico S sem especificação (amostras de tamanhos variáveis).....	15
2.2 Gráfico R	195
2.2.1 Gráfico R com especificação.....	17
2.2.2 Gráfico R sem especificação.....	17
2.3 Gráfico \bar{X}	228
2.3.1 Gráfico \bar{X} para amostras de mesmo tamanho.....	18
2.3.1.1 Gráfico \bar{X} com especificação (amostras de mesmo tamanho).....	18
2.3.1.2 Gráfico \bar{X} sem especificação (amostras de mesmo tamanho).....	19
2.3.2 Gráfico \bar{X} para amostras de tamanhos variáveis.....	20
2.3.2.1 Gráfico \bar{X} com especificação (amostras de tamanhos variáveis).....	20
2.3.2.2 Gráfico \bar{X} sem especificação (amostras de tamanhos variáveis).....	20

2.4 Exemplo 1	22
2.5 Exemplo 2	25
2.6 Exemplo 3	28
2.7 Exemplo 4	30
2.8 Considerações finais.....	33

3 Metodologia CEP para atributos – Gráficos de controle da proporção ou frequência de itens defeituosos

395

3.1 Gráfico p	36
3.1.1 Gráfico p para amostras de mesmo tamanho.....	36
3.1.1.1 Gráfico p com especificação (amostras de mesmo tamanho).....	37
3.1.1.2 Gráfico p sem especificação (amostras de mesmo tamanho).....	38
3.1.2 Gráfico p para amostras de tamanhos variáveis	39
3.1.2.1 Gráfico p com especificação (amostras de tamanhos variáveis)	40
3.1.2.2 Gráfico p sem especificação (amostras de tamanhos variáveis).....	41
3.2 Gráfico np	42
3.2.1 Gráfico np com especificação	43
3.2.2 Gráfico np sem especificação.....	43
3.3 Exemplo 5	43
3.4 Exemplo 6	46
3.5 Exemplo 7	49
3.6 Considerações finais.....	52

4 Metodologia CEP para atributos – Gráficos de controle do número de defeitos por item produzido

53

4.1 Gráfico c	53
4.2 Gráfico u	54
4.2.1 Gráfico u para amostras de mesmo tamanho.....	56

4.2.1.1	Gráfico u com especificação (amostras de mesmo tamanho).....	56
4.2.1.2	Gráfico u sem especificação (amostras de mesmo tamanho).....	56
4.2.2	Gráfico u para amostras de tamanhos variáveis	57
4.2.2.1	Gráfico u com especificação (amostras de tamanhos variáveis)	57
4.2.2.2	Gráfico u sem especificação (amostras de tamanhos variáveis).....	58
4.3	Exemplo 8	59
4.4	Exemplo 9	62
4.5	Exemplo 10	65
4.6	Considerações finais.....	68
5	Capacidade do processo	69
5.1	Índices para um processo centrado.....	71
5.2	Índices para um processo não-centrado.....	74
5.3	Exemplo 11	77
5.4	Considerações finais.....	82
6	Controle estatístico de múltiplas variáveis	84
6.1	Controle da variabilidade do processo	85
6.2	Gráfico T^2 de Hotelling	87
6.2.1	Análise com dados agrupados	88
6.2.2	Análise com dados individuais.....	91
6.3	Exemplo 12	92
6.4	Exemplo 13	96
6.5	Considerações finais.....	100
7	Capacidade do processo multivariado	101
7.1	Análise de Componentes Principais	102
7.2	Dados normais multivariados	103

7.2.1	ICPM-I	104
7.2.2	ICPM-II	106
7.2.3	Método alternativo	107
7.3	Exemplo 14	109
7.4	Considerações finais.....	114
8	Controle do processo <i>on-line</i>	116
8.1	Procedimentos <i>on-line</i> de CEP univariado.....	118
8.1.1	Entrada de dados	118
8.1.2	Confirmação das informações inseridas.....	120
8.1.3	Seleção das características para a geração dos relatórios instantâneos	121
8.1.4	Relatório instantâneo para gráficos de controle	122
8.1.5	Relatório instantâneo para a capacidade	122
8.2	Procedimentos <i>on-line</i> de CEP multivariado	123
8.2.1	Entrada de dados	125
8.2.2	Confirmação das informações inseridas.....	127
8.2.3	Seleção das características para a geração dos relatórios instantâneos	127
8.2.4	Relatório instantâneo para gráficos de controle	128
8.2.5	Relatório instantâneo para a capacidade	129
9	Conclusões e propostas futuras	132
	Referências bibliográficas	134
	Apêndice A	137
	Apêndice B	139
	Apêndice C	171

Capítulo 1

Introdução

A crescente competitividade empresarial e o aumento das exigências do consumidor, junto à globalização e à informatização mundial, têm provocado mudanças significativas na produção de ambientes manufatureiros e não-manufatureiros em todo o mundo. As empresas, em particular as indústrias, vêm, cada vez mais, a necessidade do aperfeiçoamento de seus produtos. Como consequência, o controle da qualidade de seus produtos passou a ser de extrema importância para a satisfação dos usuários dos produtos ou serviços oferecidos e para a geração de lucro.

O conceito de qualidade, de acordo com Montgomery (2004, P.6 [12]), tem seus primeiros registros antes de 1900, com Frederick W. Taylor, considerado o “Pai da Administração Científica”. Entretanto, somente em 1911, segundo Toledo (1987 [19]), o método científico de organização denominado Taylorismo foi implementado na introdução da padronização e divisão dos trabalhos, fazendo com que a montagem de produtos fosse realizada mais facilmente. Algumas características desse procedimento inicial de controle de processos são: economia de mão-de-obra, aumento da produtividade, cortes de desperdícios e segmentação das tarefas. Desde então, vários estudiosos têm aplicado seus conhecimentos na pesquisa sobre controle da produção, visando o lucro da empresa e a qualidade de seus produtos.

Existem várias formas de se definir a palavra qualidade. Para Montgomery (2004, P.2 [12]), qualidade significa a aptidão para o uso de um produto ou serviço oferecido por uma empresa. Bartmann (1986 P.1 [2]), por sua vez, reconhece: “Qualidade é um índice de satisfação que o uso do produto irá proporcionar ao consumidor”. Entretanto, a palavra qualidade tem sua primeira definição científica encontrada no Dicionário Aurélio Básico da Língua Portuguesa (1988), como: “*Propriedade, atributo ou condição das coisas ou das pessoas capaz de distinguí-las das outras e lhes determinar a natureza.*”; ou seja, a qualidade é resumida por dimensões que, juntas, determinam características aceitas nos diversos campos. Apesar da diversidade de percepções, a qualidade é um conceito de suma importância para qualquer empresa, pois é mediante a sua consideração, seu aprimoramento e sua aplicação contínua, que se pode atingir o nível de excelência em quaisquer dos objetivos fixados e atividades executadas. Um fato importante a se destacar é que, geralmente, o

conceito de qualidade é interpretado erroneamente, pois é comum as pessoas, ou até mesmo os projetistas, acreditarem que a qualidade está ligada apenas ao atendimento das especificações exigidas, quando na verdade está ligada também ao desenvolvimento do projeto do produto.

A junção da Estatística com os Procedimentos de Controle para a fiscalização da qualidade dos produtos, tendo em vista a satisfação do cliente e o lucro da empresa, foi desenvolvida em 1924 por Walter Andrew Shewhart. O papel de Shewhart foi de extrema importância nos campos da engenharia e da estatística, pois, a partir de seus estudos, foram criadas as ferramentas estatísticas mais empregadas hoje em dia para o Controle Estatístico de Processos (CEP).

A metodologia padrão para o aprimoramento da qualidade é o CEP que, segundo Montgomery (2004, P.95 [12]), consiste em um poderoso conjunto de ferramentas utilizado na obtenção da estabilidade do processo e na melhoria da capacidade através da redução da variabilidade. O CEP pode ser aplicado em qualquer tipo de processo que envolva uma sequência de passos repetitivos, isto é, pode ser aplicado em processos manufatureiros e também em processos não-manufatureiros.

Hoje em dia, existem vários programas estatísticos que podem ser usados para gerar análises sobre CEP. Contudo, a maioria desses *softwares* só pode ser usada se for adquirida sua licença, que é de custo relativamente alto para certos tamanhos e tipos de empresas. Neste trabalho é proposto um procedimento *on-line* de aplicação de algumas ferramentas estatísticas utilizadas na detecção de falhas que ocorrem em ambientes manufatureiros e não-manufatureiros. Este procedimento foi desenvolvido com base nos recursos computacionais estatísticos de um *software* livre, o *R* versão 2.5.1, e tem por objetivo viabilizar o acesso às principais ferramentas de CEP (ferramentas univariadas e multivariadas) por parte de empresas de pequeno e médio porte, principalmente.

1.1 Motivação e objetivo

Em qualquer processo, seja ele de produção de bens ou de prestação de serviços, está sempre presente a variabilidade, independente do processo ser constantemente fiscalizado ou não. Segundo Werkema (1995, P.3 [26]), essa variabilidade pode ser ocasionada devido a diferenças entre matérias-primas, às condições dos equipamentos, aos métodos de trabalho, às condições ambientais e aos operadores envolvidos no

processo. Na Estatística, o melhoramento do nível de qualidade danificado por essa variabilidade é realizado através do chamado Controle Estatístico de Processos (CEP), que consiste em um conjunto de métodos utilizados para planejar, monitorar e aprimorar um processo produtivo, através da coleta de dados de uma série de variáveis que refletem a qualidade desse processo. Quando as variáveis mensuradas no CEP são numéricas, referentes a cada produto ou serviço especificamente, diz-se que se realiza o controle de variáveis do processo; por outro lado, quando se observa a frequência de uma determinada característica (um determinado defeito de fabricação em um conjunto de produtos, por exemplo), diz-se que é realizado o controle de atributos.

Em geral, o interesse inicial do CEP está ligado ao monitoramento da variabilidade. A partir do momento em que se verifica que essa variabilidade é constante ao longo do tempo, realiza-se então o controle da média do processo. Em ambos os casos, o controle pode ser feito para observações individuais ou para observações agrupadas de acordo com alguma característica, como por exemplo, o período em que foram produzidas, a máquina ou o funcionário responsável pela produção. Posteriormente, a partir de amostras que possuem a média e a variabilidade constantes ao longo do tempo, quantifica-se a capacidade do processo de atender às especificações do cliente ou exigências impostas pelo projeto do produto. Diversas ferramentas de CEP encontram-se disponíveis, tornando viável a realização ágil e adequada do planejamento e monitoramento do processo.

Entretanto, na maioria das ocasiões em que se utilizam tais ferramentas, um fator de enorme relevância é desconsiderado: a natureza multivariada dos estudos realizados. De fato, em contextos reais, é natural que a qualidade de um processo seja refletida através de um conjunto de variáveis que normalmente exibem associação entre si. Nestes casos, a abordagem univariada da maioria das técnicas usualmente empregadas corresponde a uma limitação destas mesmas, o que resulta em conclusões menos precisas e coerentes. Em consequência disso, pode haver diminuição na satisfação e confiança do cliente, aumento nas deficiências e nos custos do processo e quedas nas receitas.

Diversos procedimentos estatísticos multivariados têm sido desenvolvidos, visando suprir a carência de técnicas que permitam a análise conjunta das variáveis de interesse. Em muitos casos, inclusive, é essencial que se disponha de metodologias que levem em consideração a estrutura de associação existente entre tais variáveis, principal

deficiência das metodologias univariadas de CEP. Evidentemente, devido ao fato de que tais metodologias são relativamente desconhecidas e recentes, exige-se que o analista disponha de amplo conhecimento técnico (estatístico) e habilidade no uso de *softwares*, dispondo de condições para, sempre que necessário, elaborar programas com as ferramentas de que necessita, tirar conclusões adequadas dos resultados observados e explicar, de modo simples, conciso e claro, para pessoas que ainda não conhecem a metodologia empregada, os detalhes que estão envolvidos na análise realizada.

Apesar da importância do CEP, sua utilização pelas indústrias de países em desenvolvimento está muito aquém do ideal, visto que os sistemas atuais exigem das companhias investimento em *software* e *hardware* para que o processo seja implantado e consolidado. Neste contexto, é evidente a vantagem de um sistema de baixo custo, que apenas necessite de um computador ligado à Internet. Este é o objetivo deste trabalho: a construção de um sistema de CEP *on-line*, de fácil utilização, que não exija investimento em *softwares* específicos, oferecendo aos possíveis usuários uma forma rápida, eficiente e precisa de utilização de procedimentos de CEP. A idéia básica consiste em prover o acesso à geração de gráficos de controle e de índices que refletem a capacidade de processos univariados e multivariados, bem como outras informações necessárias a um estudo de CEP. De uma busca minuciosa de sistemas de CEP *on-line* já disponíveis, não foi encontrado na literatura nenhum com as mesmas características do aqui desenvolvido. Assim, o sistema de CEP *on-line* proposto passa a ser, então, um recurso importante para aqueles que não têm acesso a *softwares* geralmente pagos e de alto valor. A utilização do sistema está disponível a quaisquer usuários ou analistas, ou seja, é de livre acesso.

1.2 Os dados

O uso do sistema computacional desenvolvido é exemplificado nos capítulos seguintes, a partir de dados artificiais gerados pelo *software R* versão 2.5.1. Os respectivos códigos para a geração, bem como os conjuntos de dados gerados a partir de tais códigos, estão expostos nos Apêndices A e B, respectivamente. Os exemplos de controle da qualidade são relacionados a características da qualidade de aparelhos celulares fabricados por uma suposta empresa de telecomunicações. No capítulo 2 são

expostos quatro exemplos relacionados a essa empresa fictícia, sendo que os dois primeiros estão relacionados à geração de gráficos de controle do desvio-padrão e da média (Gráficos S e \bar{X} , respectivamente), quando as amostras são de mesmo tamanho (Exemplo 1) ou de tamanhos variáveis (Exemplo 2). Os outros dois exemplos referem-se à geração de gráficos de controle da amplitude e da média (Gráficos R e \bar{X}), para dados agrupados (Exemplo 3) ou individuais (Exemplo 4). No capítulo 3 são analisados três exemplos, referentes à geração de gráficos de controle da proporção de itens defeituosos (Gráfico p), para amostras de tamanhos iguais (Exemplo 5) ou variáveis (Exemplo 6), ou da frequência de itens defeituosos produzidos (Gráfico np), quando as amostras possuem o mesmo tamanho (Exemplo 7). Três exemplos são analisados no capítulo 4, e referem-se à geração de gráficos de controle da frequência de defeitos por item produzido (Gráfico c), quando as amostras são de tamanhos iguais (Exemplo 8), ou do número médio de defeitos por item produzido (Gráfico u), quando as amostras possuem o mesmo tamanho (Exemplo 9) ou não (Exemplo 10). No capítulo 5 é exposto um único exemplo (Exemplo 11), que está relacionado à geração dos índices de capacidade do processo. O capítulo 6 expõe a análise de dois conjuntos de dados multivariados, que exemplificam a geração de gráficos de controle da variabilidade (segundo dois procedimentos estudados) e das médias (Gráfico T^2 de Hotelling), para dados agrupados (Exemplo 12) ou individuais (Exemplo 13). Por fim, no capítulo 7 é analisado um único exemplo (Exemplo 14), referente à geração de índices de capacidade do processo multivariado.

1.3 A política de controle 6-sigma

A política de controle 6-sigma é amplamente utilizada ao longo dos próximos capítulos. Visto que uma das principais atribuições do CEP é identificar mudanças que ocorrem em determinados períodos nos parâmetros do processo, são calculados dois valores de referência para cada um destes parâmetros, sendo tais valores denominados limites de controle. Para cada amostra disponível deve-se calcular uma estimativa para tais parâmetros e, a seguir, verificar se tais estatísticas se encontram entre os limites de controle. A política de controle 6-sigma determina que, dada uma especificação (ou estimativa pontual), μ , referente a determinado parâmetro (como a média, por

exemplo), e sendo conhecido (ou estimado) o desvio-padrão, σ , da estatística que estima o parâmetro em questão, os limites de controle para μ sejam dados por $\mu \pm 3\sigma$.

Essa abordagem, motivada por testes de hipóteses e inspirada na distribuição Normal, conduz a uma baixa proporção de amostras em que ocorre o erro conhecido como erro tipo I. Ou seja, é baixa a frequência de amostras que, mesmo possuindo parâmetro idêntico ao das demais amostras, são consideradas como amostras que possuem um valor discrepante para esse parâmetro. De fato, caso a distribuição da estimativa desse parâmetro seja Normal com média igual à especificação pontual, μ , e com um desvio-padrão, σ , igual ao utilizado na construção dos limites de controle, então a cada 10.000 amostras que possuem exatamente os mesmos parâmetros, 27 são consideradas como possuindo um parâmetro diferente das demais, o que corresponde a uma conclusão errônea. Dessa forma, ao se observar uma ou mais estatísticas fora dos limites de controle, é razoável verificar se houve alguma alteração nas condições de ocorrência do processo que justifique uma mudança em seus parâmetros ou se simplesmente ocorreu um alarme falso, isto é, um item foi produzido com baixo rendimento de qualidade apenas em razão da natureza aleatória da medida de qualidade mensurada.

É importante observar ainda que as metodologias a serem apresentadas são embasadas na suposição de normalidade das variáveis observadas. Portanto, é de suma importância que essa suposição seja validada antes de se tirar qualquer conclusão referente aos resultados destas metodologias. Nesse contexto, por diversas vezes é necessária a aplicação de transformações nas variáveis, de modo a obter variáveis que sigam uma distribuição Normal.

Nos próximos capítulos são discutidas algumas das ferramentas nas quais se emprega a política de controle 6-sigma. Tais ferramentas visam o monitoramento da média (ou das médias, no caso multivariado) e da variabilidade de um processo.

1.4 Organização do trabalho

Este trabalho está dividido em nove capítulos. Nos capítulos 2, 3, 4 e 5, é descrita toda a teoria de CEP utilizada no sistema *on-line*, quando apenas uma característica da qualidade do processo é observada: gráficos de controle para variáveis

(capítulo 2); gráficos de controle para atributos, separados por gráficos de controle da frequência de defeitos por item produzido (capítulo 3) e gráficos de controle do número médio de defeitos por item (capítulo 4); e teoria de capacidade (capítulo 5). Nos capítulos 6 e 7, é descrita toda a teoria de CEP utilizada no sistema *on-line*, quando duas ou mais características da qualidade do processo são observadas: gráficos de controle para a variabilidade, construídos segundo dois procedimentos distintos encontrados na literatura, e Gráfico T^2 de Hotelling (capítulo 6); e teoria de capacidade para processos multivariados (capítulo 7). No capítulo 8 é detalhado o procedimento para inserção e análise de dados de processos univariados e multivariados no sistema *on-line* desenvolvido. No capítulo 9 são apresentadas as conclusões e propostas futuras do estudo. No final do trabalho são apresentadas as referências bibliográficas estudadas e os apêndices contendo os códigos de geração dos exemplos analisados (apêndice A), os conjuntos de dados gerados a partir de tais códigos (apêndice B) e os algoritmos para construção e análise dos gráficos de controle discutidos neste trabalho (apêndice C).

Capítulo 2

Metodologia CEP para variáveis

Neste capítulo são discutidas as ferramentas estatísticas denominadas Gráficos de Controle para Variáveis, cujo objetivo é o monitoramento das variáveis que refletem a qualidade de um determinado processo, a fim de garantir que os itens resultantes desse processo apresentem um padrão de qualidade constante ao longo do tempo. O foco deste capítulo está em estudos univariados, ou seja, estudos de uma única variável ou característica da qualidade. Segundo Duncan (1986 [7]), o Dr. Walter A. Shewhart, que desenvolveu os gráficos de controle para variáveis, mais conhecidos como gráficos de Shewhart, sugeriu que os gráficos de controle devam servir, num primeiro momento, para definir a meta ou a padronização de um processo que a gestão pretenda atingir; depois, como um instrumento para alcançar essa meta; e finalmente, eles podem servir como medidas para julgar se a meta foi atingida. Esse mesmo raciocínio pode ser aplicado para os Gráficos de Controle para Atributos, que são discutidos no capítulo 3.

O interesse principal deste tipo de procedimento de controle é garantir que os parâmetros da distribuição da variável observada (como a média ou o desvio-padrão, por exemplo) não sofram efeito das variações dos diversos fatores inerentes ao processo, tais como a regulagem das máquinas, o nível de treinamento dos funcionários, a qualidade dos insumos utilizados, a organização de turnos, entre outros.

Entre as diversas ferramentas disponíveis no CEP para controle de variáveis, os gráficos de controle de Shewhart são, possivelmente, os que possuem maior destaque. Neste capítulo são discutidos os principais gráficos de Shewhart para controle de variáveis: o Gráfico S (gráfico de controle para o desvio-padrão), o Gráfico R (gráfico de controle para a amplitude) e o Gráfico \bar{X} (gráfico de controle para a média).

Um dos tipos de monitoramento da qualidade de processos é baseado na análise de desempenho de variáveis (ou características da qualidade) de natureza quantitativa, isto é, numéricas. Nesses casos, a característica em estudo pode ser representada pela variável aleatória X , com média μ e desvio-padrão σ . Exemplos desse tipo de variáveis podem ser o peso (em gramas) dos pães de alguma padaria, a quantidade (em mililitros) de suco em garrafas, o diâmetro (em metros) de parafusos de aço, a espessura

(em milímetros) de barras de chocolate produzidas por certa empresa alimentícia, entre tantos outros exemplos de aplicação de gráficos de controle para variáveis.

Ao lidar com características da qualidade de natureza quantitativa, é usualmente necessário monitorar o valor médio e a variabilidade dessas características. Geralmente, o monitoramento da média é realizado por meio do Gráfico \bar{X} e o monitoramento da variabilidade é feito através do Gráfico S ou pelo Gráfico R . Tais gráficos são gerados de forma que se tenha uma linha central, correspondente ao valor de referência (ou esperado) da característica da qualidade em estudo; um limite inferior e um limite superior de controle. Os passos para a construção desses gráficos são apresentados nas próximas seções. É fundamental realizar uma análise detalhada e conjunta, isto é, tanto para a variabilidade quanto para a média, pois pode acontecer, por exemplo, de um Gráfico \bar{X} indicar desvio na média de uma ou mais das amostras coletadas, em relação aos limites estabelecidos pela empresa ou indústria em questão, mas ao mesmo tempo apresentar a variabilidade sob controle; ou os Gráficos R ou S podem indicar que a variabilidade está fora de controle em relação aos limites de especificação ou controle, enquanto o gráfico da média apresenta estabilidade no processo.

Uma característica forte dos gráficos de controle de Shewhart é que estes são mais apropriados para detectar grandes mudanças na média ou na variabilidade de um processo. O estudo desses gráficos é realizado a partir de m amostras de tamanhos iguais ou variáveis, das quais são coletados os valores da característica X em questão (variável aleatória observada). O intuito é verificar se um determinado parâmetro, μ_X (que pode representar, por exemplo, a média da variável X), é constante ao longo de um período de tempo ou se é dependente de outros fatores intrínsecos ao processo. Para tanto, se X é a variável aleatória observada para a unidade amostral disponível e se μ_X^0 é o valor de referência ou estimativa do parâmetro μ_X , analisa-se o comportamento de X em relação a μ_X^0 e em relação aos limites de controle para μ_X , denotados por LIC (limite inferior de controle) e LSC (limite superior de controle). Essa análise é feita através de alguma estatística que represente um bom estimador do parâmetro μ_X , estatística essa capaz de agregar em um único valor toda a informação relevante disponível em cada amostra. Um processo é dito sob controle em relação ao parâmetro μ_X se os todos os valores da estatística observada se distribuírem aleatoriamente em

torno do valor de referência (linha central), μ_x^0 , e entre os limites de controle do gráfico, LIC e LSC . Valores da estatística observada que excedam os limites de controle, isto é, que estejam acima do LSC ou abaixo do LIC , existência de tendências ou sazonalidades (padrões de repetição) no gráfico construído, são sinais de falta de controle do processo.

Segundo Montgomery (2004, P.130 [12]), os gráficos de controle para a variável X podem ser usados quando a característica avaliada, X , tem distribuição Normal com média μ e desvio-padrão σ . Seja x_1, x_2, \dots, x_n uma amostra aleatória de tamanho n da variável X . A média dessa amostra pode ser descrita como:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}. \quad (2.a)$$

Como $X \sim N(\mu; \sigma)$, então:

$$\bar{X} \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right) \quad (2.b)$$

e os valores para os limites de controle são tais que atendem às exigências dadas por:

$$P[LIC \leq X \leq \mu_x^0] = \alpha/2 \text{ e } P[\mu_x^0 \leq X \leq LSC] = \alpha/2. \quad (2.c)$$

No caso de uma política de controle 6-sigma, descrita na seção 1.3, α é igual a 0,9973. Isso se deve ao fato de que, tomando esse valor para α , valores de X que se localizem fora dos limites de controle podem ser considerados muito significativos, pois, dado que α é alto, a probabilidade de erro tipo I é muito pequena. Ou seja, a probabilidade de aceitar que o processo esteja fora de controle (o parâmetro varia, muda ao longo de um período de tempo), quando na verdade está sob controle (o parâmetro é constante ao longo de todo o período de tempo considerado), é muito pequena. Além disso, caso a distribuição de X possa ser aproximada por uma Normal com média μ_x^0 e desvio-padrão σ_x , ao se assumir α igual a 0,9973, os limites de controle podem ser rescritos na forma:

$$\mu_x^0 \pm 3\sigma_x. \quad (2.d)$$

Muitas vezes é necessário estabelecer μ_x^0 e σ_x de acordo com critérios técnicos ou especificações do cliente. Quando isso não ocorre, tais parâmetros devem ser estimados a partir das amostras disponíveis.

Os gráficos de controle são usualmente analisados da seguinte forma: Gráfico S e Gráfico \bar{X} ou Gráfico R e Gráfico \bar{X} . A seguir são detalhados os gráficos de controle de Shewhart, além de sua aplicação no sistema *on-line*.

2.1 Gráfico S

O Gráfico S é usado para monitorar a variabilidade de um processo. Segundo Montgomery (2004, P.148 [12]), a variabilidade do processo pode ser monitorada adequadamente mediante a visualização do comportamento do desvio-padrão amostral, S , caso as amostras possuam pelo menos 10 ou 12 observações (isto é, quando o tamanho amostral, n , é moderadamente grande) ou caso não apresentem o mesmo tamanho (ou seja, quando n varia). Isso se deve ao fato de que, quando se dispõe de muitas observações por amostra (n moderado ou grande), o desvio-padrão amostral é mais eficiente (apresenta maior eficiência estatística) do que a amplitude amostral ao estimar a variabilidade do processo.

A seguir é descrita passo-a-passo a montagem do Gráfico S e suas possíveis maneiras de geração: quando n é constante ou varia e há ou não um valor de referência para a variabilidade do processo, representada pelo desvio-padrão σ .

2.1.1 Gráfico S para amostras de mesmo tamanho

É possível que a indústria ou empresa em questão tenha um valor de referência para a representação da variabilidade, σ . Quando isso não ocorre, tal parâmetro deve ser estimado a partir das amostras disponíveis. Nos próximos tópicos são abordadas essas duas situações no cálculo do Gráfico S , quando são coletadas amostras de mesmo tamanho: Gráfico S com especificação (seção 2.1.1.1) e Gráfico S sem especificação (seção 2.1.1.2).

2.1.1.1 Gráfico S com especificação (amostras de mesmo tamanho)

Para esse tipo de Gráfico S, a linha central é dada pelo valor de referência do desvio-padrão, σ , multiplicado por uma constante c_4 que é apresentada logo abaixo. Os limites de controle de Shewhart para o Gráfico S, obtidos a partir da expressão (2.d), são dados por:

$$\mu_s^0 \pm 3\sigma_s = c_4\sigma \pm 3\sigma\sqrt{(1-c_4^2)}. \quad (2.1.1.1.a)$$

Isso se deve ao fato de que:

$$E(S) = c_4\sigma \quad (2.1.1.1.b)$$

e

$$\sigma_s = \sigma\sqrt{(1-c_4^2)}, \quad (2.1.1.1.c)$$

com $c_4 = \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]} \sqrt{\frac{2}{n-1}}$ e $\Gamma(\cdot)$ representa a função gama. Desta forma, o valor de referência nesse gráfico deve ser igual a $c_4\sigma$ (Burr, 1976 [4]). Valores de c_4 para diversos tamanhos de amostra, além de outras constantes usadas na construção de gráficos de controle, são apresentadas na Tabela 2.1.1.1.a.

Alguns livros apresentam os limites de controle para amostras de mesmo tamanho com especificação, tal como constantes. Por exemplo, no livro do Montgomery (2004, P.149 [12]), $LIC = B_5\sigma$ e $LSC = B_6\sigma$, com:

$$B_5 = c_4 - 3\sqrt{(1-c_4^2)} \text{ e } B_6 = c_4 + 3\sqrt{(1-c_4^2)}. \quad (2.1.1.1.d)$$

Essas constantes são tabuladas para vários tamanhos amostrais na Tabela 2.1.1.1.a.

Tabela 2.1.1.1.a: Constantes usadas na construção de gráficos de controle para variáveis.

n	c_4	d_2	d_3	B_3	B_4	B_5	B_6
2	0,7979	1,128	0,853	0	3,267	0	2,606
3	0,8862	1,693	0,888	0	2,568	0	2,276
4	0,9213	2,059	0,880	0	2,266	0	2,088
5	0,9400	2,326	0,864	0	2,089	0	1,964
6	0,9515	2,534	0,848	0,030	1,970	0,029	1,874
7	0,9594	2,704	0,833	0,118	1,882	0,113	1,806
8	0,9650	2,847	0,820	0,185	1,815	0,179	1,751
9	0,9693	2,970	0,808	0,239	1,761	0,232	1,707
10	0,9727	3,078	0,797	0,284	1,716	0,276	1,669
11	0,9754	3,173	0,787	0,321	1,679	0,313	1,637
12	0,9776	3,258	0,778	0,354	1,646	0,346	1,610

13	0,9794	3,336	0,770	0,382	1,618	0,374	1,585
14	0,9810	3,407	0,763	0,406	1,594	0,399	1,563
15	0,9823	3,472	0,756	0,428	1,572	0,421	1,544
16	0,9835	3,532	0,750	0,448	1,552	0,440	1,526
17	0,9845	3,588	0,744	0,466	1,534	0,458	1,511
18	0,9854	3,640	0,739	0,482	1,518	0,475	1,496
19	0,9862	3,689	0,734	0,497	1,503	0,490	1,483
20	0,9869	3,735	0,729	0,510	1,490	0,504	1,470
21	0,9876	3,778	0,724	0,523	1,477	0,516	1,459
22	0,9882	3,819	0,720	0,534	1,466	0,528	1,448
23	0,9887	3,858	0,716	0,545	1,455	0,539	1,438
24	0,9892	3,895	0,712	0,555	1,445	0,549	1,429
25	0,9896	3,931	0,708	0,565	1,435	0,559	1,420

2.1.1.2 Gráfico S sem especificação (amostras de mesmo tamanho)

O procedimento de construção do Gráfico S sem especificação é semelhante ao descrito na seção 2.1.1.1. Contudo, quando nos referimos à não especificação, isso significa que σ , assim como a média μ , devem ser estimados a partir das amostras disponíveis. Suponha que m amostras estejam disponíveis, cada uma de tamanho n , e que S_i seja o desvio-padrão da i -ésima amostra. A média dos m desvios-padrão é:

$$\bar{S} = \frac{1}{m} \sum_{i=1}^m S_i, \quad (2.1.1.2.a)$$

com

$$S_i = \sqrt{\frac{\sum_{j=1}^n (x_{ij} - \bar{x})^2}{n-1}} \quad (2.1.1.2.b)$$

e x_{ij} é o valor da variável de interesse, X , para a observação j , $j=1,2,\dots,n$, da amostra i , $i=1,2,\dots,m$.

Um estimador não-viesado de σ , segundo Montgomery (2004, P.149 [12]), é

$$\bar{S} / c_4 \left(E \left[\frac{\bar{S}}{c_4} \right] = E \left[\frac{1}{mc_4} \sum_{i=1}^m S_i \right] = \frac{1}{mc_4} \sum_{i=1}^m E(S_i) = \sigma \right).$$

Assim, a partir de (2.d), pode-se definir que os limites de controle para o Gráfico S sem especificação e considerando amostras de mesmo tamanho, são dados por:

$$\hat{\mu}_S^0 \pm 3\hat{\sigma}_S = \bar{S} \pm 3 \frac{\bar{S}}{c_4} \sqrt{(1-c_4^2)}. \quad (2.1.1.2.c)$$

Montgomery (2004, P.149 [12]) apresenta esses limites através de constantes tabeladas que variam de acordo com o tamanho da amostra, como $LIC = B_3\bar{S}$ e $LSC = B_4\bar{S}$, com:

$$B_3 = 1 - \frac{3}{c_4} \sqrt{(1 - c_4^2)} \text{ e } B_4 = 1 + \frac{3}{c_4} \sqrt{(1 - c_4^2)}. \quad (2.1.1.2.d)$$

Valores de B_3 e B_4 para diversos tamanhos de amostra são apresentados na Tabela 2.1.1.1.a.

2.1.2 Gráfico S para amostras de tamanhos variáveis

O procedimento de construção do Gráfico S para amostras de tamanhos variáveis segue o mesmo esquema do apresentado anteriormente na seção 2.1.1. No entanto, visto que os limites de controle deste gráfico são funções do tamanho amostral n , então tais limites variam de uma amostra para outra. A seguir são abordadas duas situações que podem ocorrer no controle da qualidade utilizando o Gráfico S para amostras de tamanhos variáveis: Gráfico S com especificação (seção 2.1.2.1) e Gráfico S sem especificação (seção 2.1.2.2).

2.1.2.1 Gráfico S com especificação (amostras de tamanhos variáveis)

Em determinados casos, é possível que a indústria ou estabelecimento comercial possua valores de referência para a variabilidade do processo e que as amostras coletadas sejam de tamanhos variáveis. A diferença deste caso e do apresentado na seção 2.1.1.1 está no cálculo da constante c_4 , que é dada por:

$$c_{4i} = \frac{\Gamma(n_i / 2)}{\Gamma[(n_i - 1) / 2]} \sqrt{\frac{2}{n_i - 1}}, \quad (2.1.2.1.a)$$

$i=1,2,\dots,m$. Ou seja, é calculada uma constante c_4 para cada amostra e, conseqüentemente, um limite superior e um inferior para cada amostra, fazendo então com que os limites de controle, bem como a linha central do gráfico, não sejam mais retas.

Os limites de controle são dados então por:

$$\mu_S^0 \pm 3\sigma_S = c_{4i}\sigma \pm 3\sigma\sqrt{(1-c_{4i}^2)}. \quad (2.1.2.1.b)$$

2.1.2.2 Gráfico S sem especificação (amostras de tamanhos variáveis)

Nesse tipo de Gráfico S , é preciso fazer uma ponderação em relação ao tamanho da amostra. Suponha que m amostras estejam disponíveis, cada uma de tamanho n_i , e que \bar{S} seja a média ponderada dos m desvios-padrão. Então, segundo Montgomery (2004, P.151 [12]), tem-se que \bar{S} é dado por:

$$\bar{S} = \left[\frac{\sum_{i=1}^m (n_i - 1)S_i^2}{\left(\sum_{i=1}^m n_i \right) - m} \right]^{1/2}, \quad (2.1.2.2.a)$$

em que $S_i = \sqrt{\frac{\sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2}{n_i - 1}}$ é o desvio-padrão da i -ésima amostra.

Assim, através de (2.1.1.2.c), pode-se calcular os limites de controle para o Gráfico S sem especificação e para amostras de tamanhos variáveis. Tais limites são dados por:

$$\hat{\mu}_S^0 \pm 3\hat{\sigma}_S = \bar{S} \pm 3 \frac{\bar{S}}{c_{4i}} \sqrt{(1-c_{4i}^2)}, \quad (2.1.2.2.b)$$

sendo que c_{4i} é calculada para cada amostra através de (2.1.2.1.a).

Uma alternativa para os limites de controle variáveis, segundo Montgomery (2004, P.154 [12]), consiste em basear o cálculo dos limites de controle em um tamanho amostral médio \bar{n} , o que parece uma abordagem satisfatória caso os n_i não sejam muito diferentes, ou então no tamanho de amostra modal (mais frequente), visto que \bar{n} pode não ser um número inteiro.

2.2 Gráfico R

O Gráfico R , ou Gráfico da Amplitude Amostral, é comumente usado por engenheiros que trabalham com o CEP, por causa da sua facilidade nos cálculos para

cada amostra. Contudo, a aplicação deste tipo de gráfico é mais aconselhada nas situações em que se dispõe de amostras de mesmo tamanho e com menos de 10 ou 12 observações, em virtude de apresentar um estimador mais eficiente (que é a amplitude amostral) do que o desvio-padrão amostral para a variabilidade do processo, pois quando o valor de n é suficientemente grande existe uma perda de informação da amostra contida entre os dois valores extremos. No entanto, visto que as referências para a variabilidade do processo são, em geral, dadas em termos do desvio-padrão do processo, faz-se necessário determinar qual a relação entre o valor esperado para a amplitude amostral de cada amostra e o desvio-padrão do processo.

Então, seja a variável $R_i = \max_j(x_{ij}) - \min_j(x_{ij})$ correspondente à amplitude amostral da i -ésima amostra, isto é, a diferença entre o maior e o menor valor observados da amostra i . A relação entre a amplitude, R , de uma amostra de distribuição Normal e o desvio-padrão de tal distribuição é dada pela amplitude relativa, $W = R / \sigma$. A distribuição de W foi bem estudada, conforme Montgomery (2004, P.131 [12]), e tem como média a constante d_2 tabelada ($E(W) = d_2$), cujo valor depende do tamanho da amostra. Valores da constante d_2 para vários tamanhos amostrais podem ser vistos na Tabela 2.1.1.1.a. Portanto, um estimador não-viesado de σ é:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}, \quad (2.2.a)$$

$$\text{com } \bar{R} = \frac{1}{m} \sum_{i=1}^m R_i \left(E \left[\frac{\bar{R}}{d_2} \right] = E \left[\frac{1}{md_2} \sum_{i=1}^m R_i \right] = \frac{1}{md_2} \sum_{i=1}^m E(R_i) = \sigma \right).$$

Neste trabalho é abordado somente o uso do Gráfico R para tamanhos iguais de amostra, visto que o uso deste gráfico para amostras de tamanhos variáveis não é indicado devido à sua complexidade (pois tanto os limites de controle quanto a linha de referência do Gráfico R variam quando n é não-constante). Nos tópicos seguintes são abordadas a construção do Gráfico R quando se dispõe (seção 2.2.1) ou não (seção 2.2.2) de uma especificação (ou referência) para o desvio-padrão σ da variável observada.

2.2.1 Gráfico R com especificação

Considerando W , como já foi visto na seção 2.2, tem-se que $E(W) = E(R/\sigma) = d_2$. Portanto, considerando que σ é uma constante, tem-se que $E(R) = d_2\sigma$, que deve ser usado como valor de referência do Gráfico R . Para o cálculo dos limites de controle deste gráfico, é necessário, primeiro, obter o desvio-padrão da amplitude amostral, σ_R . E, para isso, é preciso estabelecer o valor da constante $d_3 = \sqrt{\text{Var}(R/\sigma)}$, pois o que se quer é $\sigma_R = \sqrt{\text{Var}(R)} = d_3\sigma$. Assim, tem-se que os limites de controle, a partir da equação de limites base em (2.d), são dados por:

$$\mu_R^0 \pm 3\sigma_R = d_2\sigma \pm 3d_3\sigma. \quad (2.2.1.a)$$

Os valores de d_2 e d_3 variam de acordo com o tamanho da amostra e estão expostos na Tabela 2.1.1.1.a.

2.2.2 Gráfico R sem especificação

Quando a construção do Gráfico R não contém referências (ou especificações), deve-se estimar o parâmetro em questão, σ . Nesse caso, deve-se estimar o valor de referência do Gráfico R por:

$$\bar{R} = \frac{1}{m} \sum_{i=1}^m R_i, \quad (2.2.2.a)$$

visto que esse é o estimador da média das amplitudes amostrais, com $R_i = \max_j(x_{ij}) - \min_j(x_{ij})$.

Assim como quando existe especificação, tem-se neste caso, que a partir de (2.2.a) e de $\sigma_R = \sqrt{\text{Var}(R_i)} = d_3\sigma$, o estimador para o desvio-padrão da amplitude amostral é dado por $\hat{\sigma}_R = \frac{d_3\bar{R}}{d_2}$. Portanto, os limites de controle do Gráfico R são dados

por:

$$\hat{\mu}_R^0 \pm 3\hat{\sigma}_R = \bar{R} \pm 3\frac{d_3\bar{R}}{d_2}. \quad (2.2.2.b)$$

O Gráfico R com ou sem especificação pode ser aplicado também quando se observam amostras de tamanho unitário ($n=1$). Nesse caso, substituem-se as amplitudes amostrais, R_i , pelas amplitudes móveis, dadas por $R_i = |x_i - x_{i-1}|$, $i=1,2,\dots,m$.

2.3 Gráfico \bar{X}

O controle da média do processo é usualmente feito através do Gráfico \bar{X} . Tal gráfico pode ser construído segundo diversas condições que são detalhadas nesta seção. As maneiras diferentes de se construir este gráfico são: com especificação, sem especificação, com mesmo tamanho de amostra e com amostras de tamanhos variáveis.

2.3.1 Gráfico \bar{X} para amostras de mesmo tamanho

A construção do Gráfico \bar{X} para amostras de mesmo tamanho baseia-se na montagem de três retas, sendo o valor especificado ou estimado da média como a reta central (linha central ou de referência), e as demais retas representando os limites superior (LSC) e inferior de controle (LIC) do gráfico. É possível que a indústria ou o estabelecimento comercial em questão possua um valor de referência (ou especificação) para a média do processo. Devido a isso, são discutidas a seguir as duas maneiras de construção do Gráfico \bar{X} para amostras de mesmo tamanho: com especificação (seção 2.3.1.1) e sem especificação (seção 2.3.1.2).

2.3.1.1 Gráfico \bar{X} com especificação (amostras de mesmo tamanho)

A palavra *especificação* indica, neste caso, que o valor do parâmetro em questão, isto é, a média, foi definida pela política do local ou por órgãos governamentais, como o INMETRO, por exemplo. Assim, neste gráfico pode-se verificar a constância do parâmetro (a média do processo) ao longo do tempo. O Gráfico \bar{X} apresenta três linhas ou retas: uma central, correspondente à especificação μ^0 , no caso, e as demais representando os limites de controle dados em (2.d) que, para serem calculados, é

preciso obter o valor do desvio-padrão da média amostral, que sob a suposição de independência entre as observações, é dado por $\sigma_{\bar{x}} = \sigma / \sqrt{n}$ (veja a equação (2.b)) e que, por sua vez, necessita do valor de σ , que pode ter sido especificado ou não.

Assim, obtêm-se os limites de controle para a média do processo, dados por:

$$\mu_{\bar{x}}^0 \pm 3\sigma_{\bar{x}} = \mu_x^0 \pm 3 \frac{\sigma_x}{\sqrt{n}} = \mu^0 \pm 3 \frac{\sigma}{\sqrt{n}}. \quad (2.3.1.1.a)$$

2.3.1.2 Gráfico \bar{X} sem especificação (amostras de mesmo tamanho)

No contexto em que se deseja monitorar um processo, mas não se dispõe de referências (ou especificações) para os parâmetros da distribuição da variável observada, faz-se necessário estimá-los a partir das amostras disponíveis. Portanto, é preciso encontrar um estimador para μ^0 que represente a média da variável observada.

Sendo \bar{x}_i a média para a variável em questão, considerando a i -ésima amostra, dentre as m amostras disponíveis, então um possível estimador para μ^0 representando a média, é a média das médias ou média geral, dada por:

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m}. \quad (2.3.1.2.a)$$

Isso é fato, pois $E(\bar{\bar{x}}) = E\left(\frac{1}{m} \sum_{i=1}^m \bar{x}_i\right) = E(\bar{x}_i) = E\left(\frac{1}{n} \sum_{j=1}^n x_{ij}\right) = E(x_{ij}) = \mu^0$, desde que o processo se encontre sob controle, o que justifica seu uso como valor de referência (ou linha central) do Gráfico \bar{X} .

Um estimador não-viesado do desvio-padrão da média de um processo sob controle, que considera todas as amostras disponíveis, é dado por:

$$\hat{\sigma}_{\bar{x}} = \frac{\bar{S}}{c_4 \sqrt{n}}. \quad (2.3.1.2.b)$$

Esse resultado é válido, pois partindo da equação referente ao desvio-padrão de X ($E(S_i) = c_4 \sigma$), tem-se que $E\left[\frac{\bar{S}}{c_4}\right] = E\left[\frac{1}{mc_4} \sum_{i=1}^m S_i\right] = \sigma$, desde que o processo esteja sob controle.

Assim, pode-se calcular os limites de controle do Gráfico \bar{X} sem especificação e para amostras de mesmo tamanho, através de:

$$\hat{\mu}_{\bar{x}}^0 \pm 3\hat{\sigma}_{\bar{x}} = \bar{\bar{x}} \pm 3 \frac{\bar{S}}{c_4 \sqrt{n}}. \quad (2.3.1.2.c)$$

2.3.2 Gráfico \bar{X} para amostras de tamanhos variáveis

Em determinados processos, industriais e não-industriais, eventualmente não se têm tamanhos iguais de amostra, o que pode provocar algumas mudanças nos cálculos do Gráfico \bar{X} . A seguir são apresentados esses cálculos quando coletadas m amostras de tamanhos variáveis.

2.3.2.1 Gráfico \bar{X} com especificação (amostras de tamanhos variáveis)

Quando a indústria ou empresa em que está sendo realizado o CEP fornece uma referência (ou especificação) para a média do processo, então não é necessário estimar μ^0 . Neste caso, a linha central do Gráfico \bar{X} é dada pela especificação μ^0 e os limites de controle podem ser calculados considerando-se a equação (2.d), com a diferença de que agora os limites devem ser calculados para cada amostra diferentemente, pois os n 's variam. Assim, o cálculo dos limites de controle é dado por:

$$\mu^0 \pm 3 \frac{\sigma}{\sqrt{n_i}}. \quad (2.3.2.1.a)$$

2.3.2.2 Gráfico \bar{X} sem especificação (amostras de tamanhos variáveis)

Em determinadas situações não é possível obter especificações para monitorar a qualidade de um processo. Nesses casos, como não há exigências em relação ao parâmetro, no caso a média do processo, é necessário então estimá-la a partir das amostras disponíveis.

O valor de referência do Gráfico \bar{X} é dado pela estimativa do parâmetro, uma média ponderada das médias \bar{x}_i , que é dada por:

$$\bar{\bar{x}} = \frac{\sum_{i=1}^m n_i \bar{x}_i}{\sum_{i=1}^m n_i} = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^m n_i}. \quad (2.3.2.2.a)$$

Essa alteração (a ponderação no cálculo da média) visa garantir que todas as observações tenham o mesmo peso sobre a estimativa da média do processo. Observe ainda que, devido a essa alteração, $\bar{\bar{x}}$ continua sendo um estimador não-viesado de μ^0 ,

$$\text{pois } E(\bar{\bar{x}}) = E\left(\frac{\sum_{i=1}^m n_i \bar{x}_i}{\sum_{i=1}^m n_i}\right) = E\left(\frac{\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^m n_i}\right) = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} E(x_{ij})}{\sum_{i=1}^m n_i} = \frac{\sum_{i=1}^m n_i E(x_{ij})}{\sum_{i=1}^m n_i} = E(x_{ij}) = \mu^0.$$

Da equação (2.3.1.2.b), tem-se que o valor do desvio-padrão da média, quando os parâmetros não são especificados e o tamanho das amostras varia, é $\frac{\bar{S}}{c_{4i}\sqrt{n_i}}$, com

$$c_{4i} = \frac{\Gamma(n_i/2)}{\Gamma[(n_i-1)/2]} \sqrt{\frac{2}{n_i-1}}, \quad i=1,2,\dots,m.$$

Assim, os limites de controle são dados por:

$$\bar{\bar{x}} \pm 3 \frac{\bar{S}}{c_{4i}\sqrt{n_i}}. \quad (2.3.2.2.b)$$

Para os gráficos discutidos neste capítulo, vale ressaltar que apenas as amostras que não sofreram influência de causas (ou desajustes) eventualmente observadas devem ser utilizadas na estimação dos limites de controle dos gráficos e dos parâmetros do processo, caso exista interesse no uso, em análises posteriores com outras amostras, dos limites de controle estimados. Todas as amostras que sofreram influência de tais causas, independentemente de estarem sob controle ou não, devem ser desconsideradas na estimação dos limites de controle.

Algoritmos para construção e análise dos gráficos de controle para variáveis são apresentado nos Quadros 2.1.a e 2.2.a do Apêndice C.

2.4 Exemplo 1

O primeiro conjunto de dados analisado neste capítulo, cujos valores estão na Tabela 2.4 do Apêndice B, contém 25 amostras com 15 observações cada, representando a capacidade de armazenamento (em megabytes, MB) de cartões de memória para aparelhos celulares de uma marca qualquer (dados artificiais). Suponha que o interesse da empresa é manter o processo sob controle, com especificações (valores de referência) de 100 MB para a média e desvio-padrão igual a 10 MB.

A seguir, nas Figuras 2.4.a, 2.4.b, 2.4.c e 2.4.d é apresentada uma sequência de passos para a análise desse conjunto de dados no sistema de CEP *on-line* desenvolvido, desde a inserção dos dados (Figura 2.4.a) até a visualização do relatório instantâneo para os gráficos de controle S e \bar{X} (Figura 2.4.d).

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Peso
93.26
104.76
100.82
105.82
104.63
112.62
105.47
90.99
92.10
86.53
93.65
105.68
```

Figura 2.4.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo1
Número de variáveis:	1
Número de observações:	375
Número de amostras:	25
Tamanho das amostras:	15

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 2.4.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu: Variáveis

Escolha os gráficos adequados para seu tipo de dados: Gráficos S e Xbarra

Defina as condições para cada gráfico: Com especificação

Média: 100

Desvio-Padrão: 10

Avançar

Figura 2.4.c: Definindo as características e condições do processo.

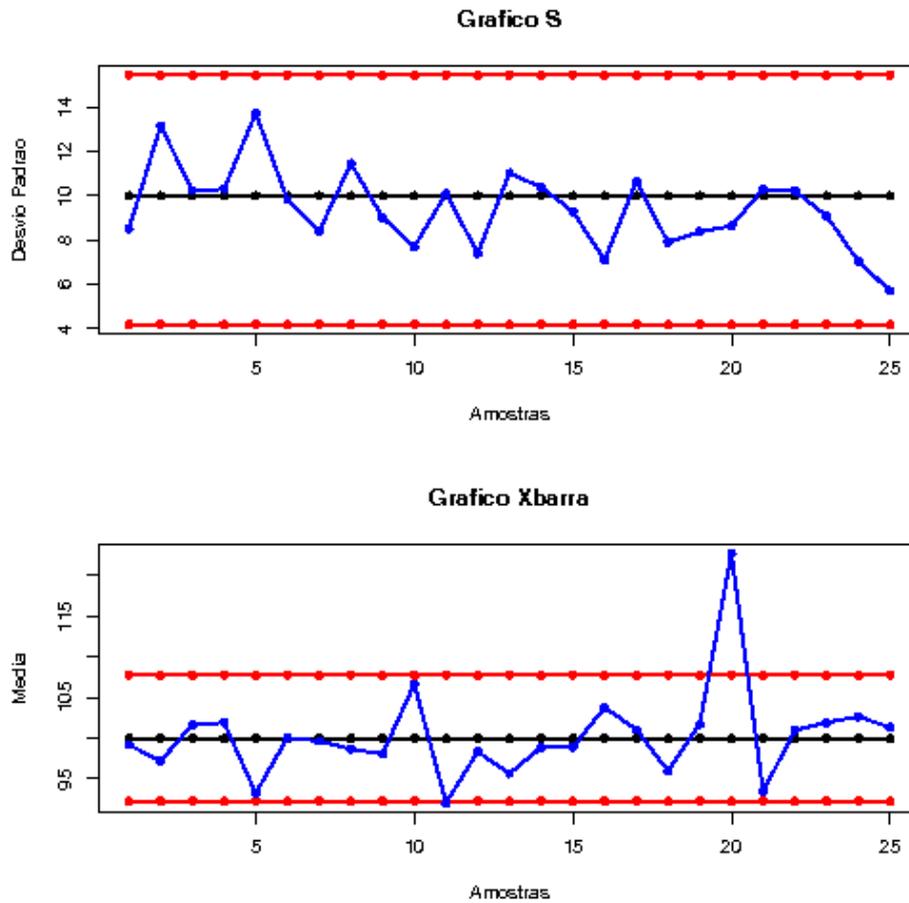


Figura 2.4.d: Relatório instantâneo gerado para os Gráficos S e \bar{X} .

O Gráfico S apresentado na Figura 2.5.d não detectou qualquer sinal de falta de controle do processo em questão, com respeito à variabilidade. No entanto, o Gráfico \bar{X} , apresentado nessa mesma figura, identificou um sinal de falta de controle pontual do processo, referente à amostra de número 20, cuja causa deve ser investigada. Continuando com a análise (que não é feita aqui), esta amostra de número 20 deve ser retirada e os gráficos de controle refeitos, para saber se há mais amostras fora de controle ou se realmente ocorreu uma falta de controle pontual do processo. Para o caso em que os parâmetros do processo, isto é, a média e o desvio-padrão, são estimados, o comportamento dos gráficos (que também não são exibidos aqui) é equivalente ao observado naqueles presentes na Figura 2.5.d.

2.5 Exemplo 2

No segundo conjunto de dados analisado, cujos valores estão na Tabela 2.5 do Apêndice B, também é observada a capacidade de armazenamento (em megabytes, MB) de cartões de memória para aparelhos celulares. Contudo, as 25 amostras coletadas são de tamanhos desiguais que variam entre 14 e 16 cartões. As especificações (valores de referência) para esse processo ainda se mantêm em uma média de 100 MB e um desvio-padrão de 10 MB.

A seguir, nas Figuras 2.5.a, 2.5.b, 2.5.c e 2.5.d são apresentados os passos necessários para a análise desse conjunto de dados no sistema de CEP *on-line*, desde a inserção dos dados (Figura 2.5.a) até a visualização do relatório instantâneo para os gráficos de controle S e \bar{X} (Figura 2.5.d).

Assim como no exemplo anterior, o Gráfico S apresentado na Figura 2.5.d não detectou pontos fora dos limites de controle. Note, no entanto, que as 5 últimas amostras se encontram acima da linha central do gráfico, o que pode ser um eventual sinal de falta de controle do processo com relação à variabilidade. O Gráfico \bar{X} , por sua vez, identificou um sinal de falta de controle pontual do processo, referente à amostra de número 20, que se encontra acima do limite superior deste gráfico. Identificada uma causa como responsável pela falta de controle do processo, a análise (que não é realizada aqui) prossegue, retirando esta amostra de número 20 do banco de dados e reconstruindo os gráficos de controle a partir das demais amostras, para saber se há mais pontos fora de controle ou se realmente ocorreu uma falta de controle pontual do processo. Para o caso em que os parâmetros do processo (média e desvio-padrão) são estimados, o comportamento dos gráficos (que também não são exibidos aqui) é equivalente ao observado naqueles presentes na Figura 2.5.d.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

```
16 14 15 16 15 14 15 15 16
16 15 15 15 15 14 15 16 16
15 15 16 16 16 15 14
```

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Peso
104.26
102.09
103.13
100.08
105.61
114.14
93.78
91.30
101.29
99.49
94.98
96.47
```

Figura 2.5.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo2
Número de variáveis:	1
Número de observações:	380
Número de amostras:	25
Tamanho das amostras:	vários

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 2.5.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Média:

Desvio-Padrão:

Figura 2.5.c: Definindo as características e condições do processo.

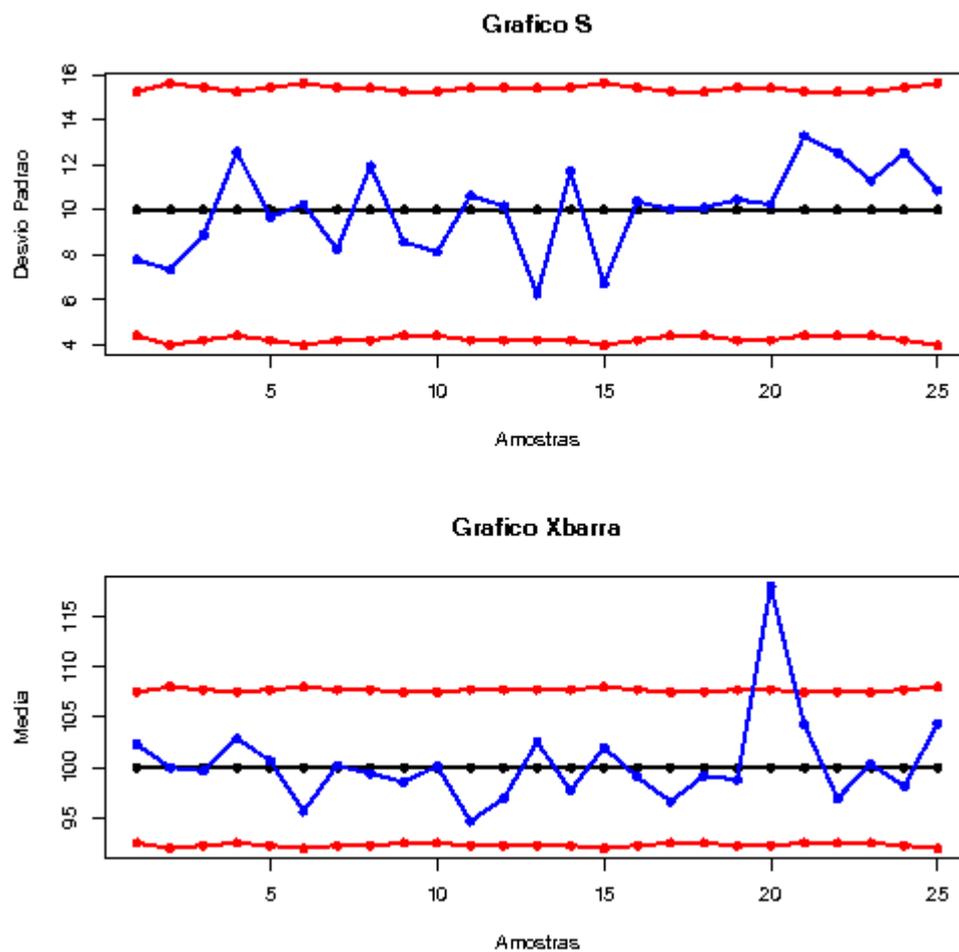


Figura 2.5.d: Relatório instantâneo gerado para os Gráficos S e \bar{X} .

2.6 Exemplo 3

O terceiro conjunto de dados analisado, cujos valores estão na Tabela 2.6 do Apêndice B, refere-se às amostras obtidas da mesma linha de produção (de cartões de memória) que os dois primeiros (Exemplos 1 e 2), em que são observadas as capacidades de armazenamento (em megabytes, MB) de cartões de memória para aparelhos celulares. Tal conjunto é constituído de 25 amostras de tamanhos iguais a 4. Suponha que o interesse da empresa de telecomunicações seja manter o processo sob controle, mas esta não tem especificações (referências) em relação aos parâmetros (média e desvio-padrão).

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Peso
110.61
95.00
108.43
95.30
104.98
99.57
107.89
101.77
98.19
112.81
89.93
90.88
```

Figura 2.6.a: Inserindo os dados.

Nas Figuras 2.6.a, 2.6.b, 2.6.c e 2.6.d é apresentada uma sequência de passos para a análise desse conjunto de dados no sistema de CEP *on-line* proposto, desde a inserção dos dados (Figura 2.6.a) até a visualização do relatório instantâneo para os gráficos de controle R e \bar{X} (Figura 2.6.d).

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo3
Número de variáveis:	1
Número de observações:	100
Número de amostras:	25
Tamanho das amostras:	4

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 2.6.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu: Variáveis

Escolha os gráficos adequados para seu tipo de dados: Gráficos R e Xbarra

Defina as condições para cada gráfico: Sem especificação

Avançar

Figura 2.6.c: Definindo as características e condições do processo.

O Gráfico R apresentado na Figura 2.6.d não detectou qualquer sinal de falta de controle do processo, com respeito à variabilidade. No entanto, o Gráfico \bar{X} , apresentado nessa mesma figura, identificou um sinal de falta de controle pontual do processo, referente à amostra de número 20, cuja causa deve ser investigada.

Continuando com a análise (que não é realizada aqui), esta amostra de número 20 deve ser retirada e os gráficos de controle refeitos, para saber se há mais amostras fora de controle ou se realmente ocorreu uma falta de controle pontual do processo.

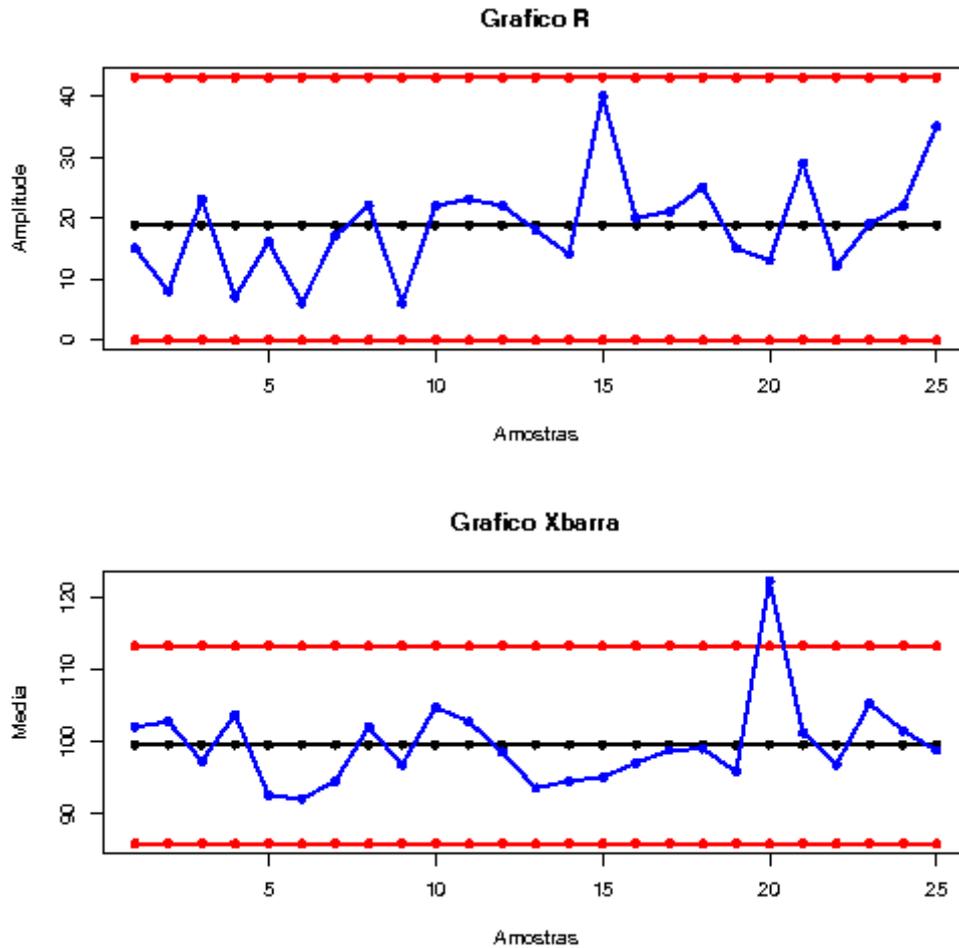


Figura 2.6.d: Relatório instantâneo gerado para os Gráficos R e \bar{X} .

2.7 Exemplo 4

O último conjunto de dados analisado neste capítulo, cujos valores estão na Tabela 2.7 do Apêndice B, contém 25 amostras de tamanhos iguais a um, coletadas da mesma linha de produção de cartões de memória dos exemplos anteriores. A característica de qualidade observada ainda é a capacidade de armazenamento (em megabytes, MB) dos cartões fabricados. Suponha que o interesse da empresa é manter o

processo sob controle, considerando as especificações (valores de referência) de 100 MB para a média e de 10 MB para o desvio-padrão.

A seguir, nas Figuras 2.7.a, 2.7.b, 2.7.c e 2.7.d é apresentada uma sequência de passos para a análise do quarto conjunto de dados no sistema de CEP *on-line*, desde a inserção dos dados (Figura 2.7.a) até a visualização do relatório instantâneo para os gráficos de controle R e \bar{X} (Figura 2.7.d).

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Peso
99.44
106.63
85.78
101.67
110.93
122.88
116.47
87.20
101.49
109.06
105.51
108.19
```

Figura 2.7.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo4
Número de variáveis:	1
Número de observações:	25
Número de amostras:	25
Tamanho das amostras:	1

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 2.7.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu: Variáveis

Escolha os gráficos adequados para seu tipo de dados: Gráficos R e Xbarra

Defina as condições para cada gráfico: Com especificação

Média: 100

Desvio-Padrão: 10

Avançar

Figura 2.7.c: Definindo as características e condições do processo.

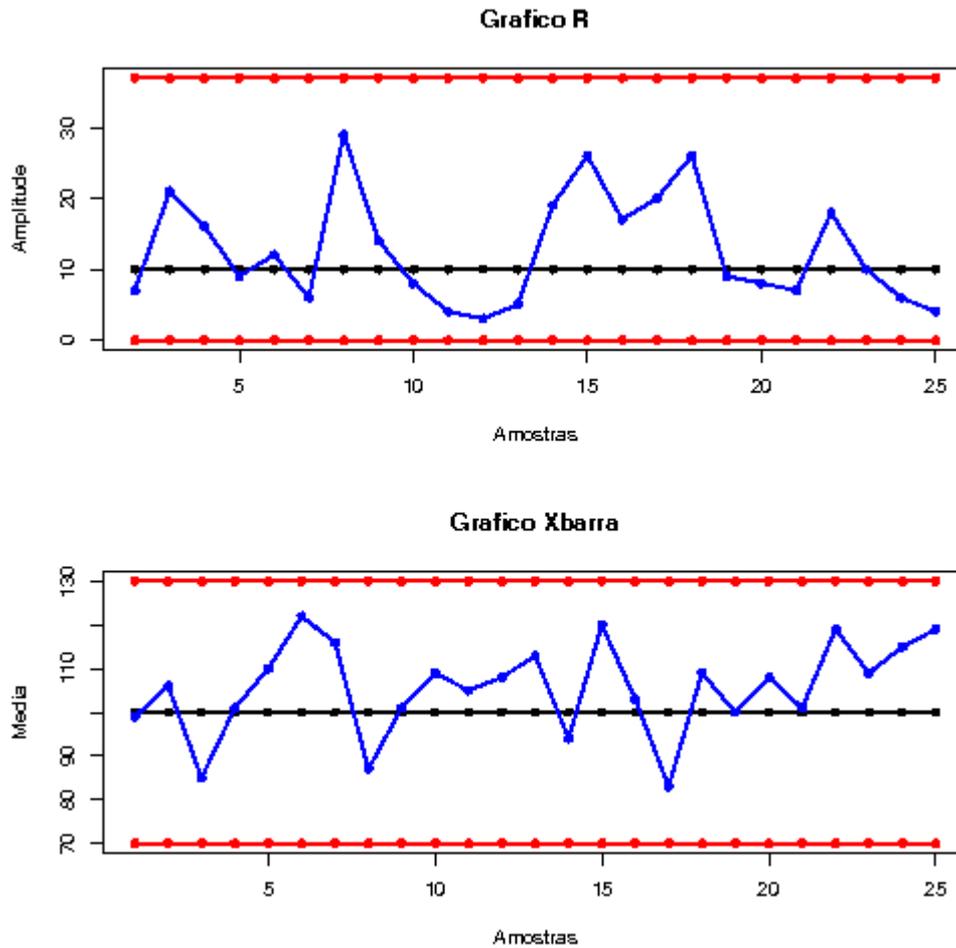


Figura 2.7.d: Relatório instantâneo gerado para os Gráficos R e \bar{X} .

Como pode ser visto na Figura 2.7.d, não há pontos fora dos limites de controle nem no Gráfico R nem no Gráfico \bar{X} . No entanto, são observados alguns pontos consecutivos acima da linha central em ambos os gráficos. O comportamento de tais gráficos para o caso em que não são consideradas as especificações dos parâmetros (tais gráficos não são mostrados aqui) é equivalente ao observado na Figura 2.7.d.

2.8 Considerações finais

Foram discutidos neste capítulo os métodos básicos do controle estatístico de processos (CEP), como os gráficos de controle de Shewhart, bem como exemplos de sua aplicação no sistema de CEP *on-line*. Ou seja, discutiu-se a aplicação do CEP no

monitoramento de processos nos quais o nível de qualidade dos itens produzidos é refletido por uma variável quantitativa. Contudo, em muitas empresas, a qualidade dos itens produzidos está relacionada a critérios não quantitativos. Pode ser que o item seja simplesmente classificado como adequado ou inadequado para comercialização (ou como em bom estado ou defeituoso). Em outros casos, pode ser que seja observado se existem um ou mais tipos de defeitos presentes no item produzido. Nestes casos, o monitoramento realizado é denominado Controle de Atributos e a sua abordagem é diferenciada da realizada até então. Detalhes a respeito da análise de controle de atributos são apresentados nos capítulos 3 e 4, junto com exemplos de sua aplicação no sistema de CEP *on-line* desenvolvido.

Capítulo 3

Metodologia CEP para atributos – Gráficos de controle da proporção ou frequência de itens defeituosos

O CEP é bastante empregado para o controle de variáveis com características mensuráveis, isto é, quantitativas. Entretanto, existem situações nas quais as características da qualidade de um processo são representadas de maneira não-numérica, ou seja, através de dados qualitativos. Geralmente, as observações dessas variáveis qualitativas são expressas por defeituoso e não-defeituoso (em bom estado), ou por conforme e não-conforme. Um exemplo de uso desse tipo de variável pode ser aplicado em uma empresa que produz canetas e deseja verificar se estas são capazes de desempenhar sua função logo que retiradas da embalagem, isto é, o interesse reside em saber se as canetas emitem tinta mediante a tentativa de escrita.

Contudo, embora as variáveis observadas inicialmente sejam qualitativas, estas não são as monitoradas na prática pelo CEP. De fato, quando se observa se as características para um determinado item se enquadram em seu projeto, também se observam automaticamente duas variáveis quantitativas que também refletem a qualidade do processo: a frequência com que se produzem itens defeituosos e o número de defeitos observados em uma determinada quantidade de itens produzidos. O controle destas variáveis é usualmente denominado Controle de Atributos.

A motivação dos gráficos de controle para atributos é a mesma dos gráficos de controle de Shewhart, discutida no capítulo 2. Em resumo, deve-se construir o gráfico temporal da estatística de interesse. Essa estatística corresponde a um estimador de um determinado parâmetro, μ , que pode corresponder, por exemplo, ao número esperado de defeitos em uma amostra com n unidades. São então representados no gráfico os limites do intervalo de confiança para μ (denominados limites de controle), bem como uma estimativa pontual ou o valor ideal (especificado) para μ . Caso a série temporal não se distribua aleatoriamente em torno da estimativa pontual (linha central do gráfico), apresente alguma tendência ou sazonalidade (padrão de repetição), ou não esteja restrita ao intervalo estabelecido pelos limites de controle, pode-se concluir que provavelmente houve, por um determinado período de tempo, alguma alteração nas condições do processo que levou à mudança no valor do parâmetro μ , isto é, à falta de

controle do processo. Dessa forma, faz-se necessário, primariamente, investigar quais as alterações que possivelmente levaram a essa mudança. Em seguida, deve-se estabelecer medidas estratégicas que evitem ou minimizem tais mudanças ou políticas gerenciais que neutralizem os efeitos negativos dessa mudança para a empresa.

Os gráficos de controle da proporção ou frequência de itens defeituosos são úteis especialmente quando um item pode ser classificado como adequado ou inadequado, em bom estado ou defeituoso. Neste caso, há interesse em garantir que a frequência com que se produzem itens incapazes de cumprir adequadamente com suas atribuições seja baixa e com um valor esperado constante ao longo de um período de tempo. Os principais gráficos de controle desse tipo são: o gráfico de controle da proporção de itens defeituosos, ou Gráfico p , e o gráfico de controle da frequência de itens defeituosos, ou Gráfico np , discutidos em detalhes nas seções que se seguem.

3.1 Gráfico p

O gráfico de controle da proporção de itens defeituosos, denominado Gráfico p , representa, na forma de uma série temporal, as proporções de itens defeituosos observados nas amostras coletadas. Esse gráfico é útil quando se deseja monitorar a probabilidade de um item qualquer não satisfazer às especificações. Naturalmente, quando essa probabilidade for elevada, haverá um destacado número médio de itens defeituosos produzidos. Por outro lado, se essa probabilidade for demasiadamente baixa, é possível que os custos decorrentes dos esforços em se melhorar o nível de precisão na linha de produção sejam altos demais.

Os gráficos para atributos, assim como os gráficos para variáveis vistos no capítulo 2, também podem ser gerados com ou sem especificações. Nos próximos tópicos são abordadas estas duas situações na construção do Gráfico p quando há amostras de mesmo tamanho ou amostras de tamanhos variáveis.

3.1.1 Gráfico p para amostras de mesmo tamanho

Nesta seção são abordados os usos do Gráfico p em amostras de mesmo tamanho. Os dois tópicos estudados a seguir são: a construção do Gráfico p com

especificação (seção 3.1.1.1) e a construção do Gráfico p sem especificação (seção 3.1.1.2).

3.1.1.1 Gráfico p com especificação (amostras de mesmo tamanho)

Em determinadas indústrias ou locais em que se deseja saber sobre a qualidade de seus produtos ou serviços oferecidos, muitas vezes são estabelecidas especificações que determinam qual deve ser a probabilidade p de que um item qualquer produzido seja defeituoso, para viabilizar financeiramente o processo e evitar a produção de um número alto de itens defeituosos.

O Gráfico p com especificação é construído da mesma maneira que os gráficos de controle apresentados no capítulo 2: existe uma linha central representando o valor de referência (especificação para a probabilidade p) e duas linhas representando os limites de controle (LIC e LSC). Para a construção desses limites, é preciso determinar, a partir de informações conhecidas, qual o valor esperado e o desvio-padrão da estatística \hat{p} (proporção amostral). Montgomery (2004, P.178 [12]) afirma que o Gráfico p se baseia na distribuição binomial, pois supondo um processo com unidades produzidas independentemente e que p seja a probabilidade de que alguma unidade seja defeituosa, então se houver n unidades do produto e D for o número de defeituosas, tem-se que $D \sim Binomial(n, p)$. Assim, supondo que sejam coletadas m amostras com n observações cada, a proporção de unidades defeituosas para a i -ésima amostra, é dada por:

$$\hat{p}_i = \frac{1}{n} \sum_{j=1}^n x_{ij} = \frac{D_i}{n}, \quad (3.1.1.1.a)$$

tal que x_{ij} assume os valores:

$$x_{ij} = \begin{cases} 1, & \text{se o } j\text{-ésimo item da } i\text{-ésima amostra for defeituoso} \\ 0, & \text{c.c.} \end{cases}, \quad (3.1.1.1.b)$$

com $P(x_{ij} = 1) = p$ e $P(x_{ij} = 0) = 1 - p = q$.

O valor esperado para a proporção amostral de defeitos \hat{p}_i , é igual à probabilidade p de que cada unidade pertencente a essa amostra seja defeituosa. Para demonstrar isso, note que:

$$E(\hat{p}_i) = E\left(\frac{1}{n} \sum_{j=1}^n x_{ij}\right) = \frac{1}{n} \sum_{j=1}^n E(x_{ij}) = E(x_{ij}). \quad (3.1.1.1.c)$$

Como x_{ij} , definido em (3.1.1.1.b), é uma variável com distribuição de Bernoulli com parâmetro p , então a equação (3.1.1.1.c) resulta em:

$$E(\hat{p}_i) = p. \quad (3.1.1.1.d)$$

O desvio-padrão de \hat{p}_i é dado por:

$$DP(\hat{p}_i) = \sqrt{\frac{p(1-p)}{n}}, \quad (3.1.1.1.e)$$

pois,

$$DP(\hat{p}_i) = \sqrt{\text{Var}(\hat{p}_i)} = \sqrt{\text{Var}\left(\frac{1}{n} \sum_{j=1}^n x_{ij}\right)} = \sqrt{\frac{1}{n^2} \text{Var}\left(\sum_{j=1}^n x_{ij}\right)}. \quad (3.1.1.1.f)$$

Devido à suposição de independência entre as observações pertencentes à mesma amostra e ao fato de que x_{ij} segue uma distribuição de Bernoulli com parâmetro p , conclui-se, a partir da equação (3.1.1.1.f) que a equação (3.1.1.1.e) é verdadeira, pois, neste caso,

$$DP(\hat{p}_i) = \sqrt{\frac{1}{n^2} \sum_{j=1}^n \text{Var}(x_{ij})} = \sqrt{\frac{np(1-p)}{n^2}} = \sqrt{\frac{p(1-p)}{n}}. \quad (3.1.1.1.g)$$

Quando o processo está sob controle, a probabilidade de ocorrência de um defeito é constante para todos os itens produzidos, independentemente da amostra à qual estes pertencem. Dessa forma, se essa probabilidade for especificada como sendo igual a p , o valor de referência (linha central) do gráfico de controle da proporção de itens defeituosos deve ser dado por p , visto que esse é o valor esperado para essa proporção em cada amostra. Com base nas equações (3.1.1.1.d) e (3.1.1.1.e), são construídos os limites de controle 6-sigma desse gráfico, definidos por:

$$p \pm 3\sqrt{\frac{p(1-p)}{n}}. \quad (3.1.1.1.h)$$

3.1.1.2 Gráfico p sem especificação (amostras de mesmo tamanho)

Quando não existe especificação para o parâmetro p , é preciso estimá-lo a partir de informações de dados existentes, isto é, a partir das amostras disponíveis. Se forem

consideradas m amostras de tamanho n cada, calcula-se $\hat{p}_i = \frac{D_i}{n}$ para cada amostra i , $i=1, \dots, m$, e a linha de referência (ou linha central) do Gráfico p é a estimação da probabilidade de que um item qualquer produzido seja defeituoso, dada por:

$$\tilde{p} = \bar{p} = \frac{\sum_{i=1}^m D_i}{mn} = \frac{1}{m} \sum_{i=1}^m \hat{p}_i. \quad (3.1.1.2.a)$$

Note que \bar{p} é um estimador não-viesado para p , pois

$$E(\bar{p}) = E\left(\frac{1}{m} \sum_{i=1}^m \hat{p}_i\right) = \frac{1}{m} \sum_{i=1}^m E(\hat{p}_i) = E(\hat{p}_i) = p.$$

Assim, de acordo com as equações (3.1.1.1.h) e (3.1.1.2.a), tem-se que os limites de controle são dados por:

$$\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}. \quad (3.1.1.2.b)$$

Um ponto a se destacar é que um processo pode estar fora de controle em relação às especificações estabelecidas para o processo, sem estar fora de controle em relação aos parâmetros estimados a partir das amostras disponíveis. Isso significa que, embora o processo não esteja apto a atender às exigências preestabelecidas, não há efeito de causas especiais que conduzam a um comportamento atípico para o processo. O que ocorre, na verdade, é que se faz necessário rever o planejamento desse processo, de tal forma a melhorar sua capacidade de produzir itens dentro dos padrões de qualidade desejados.

No contexto em que não há especificações para o processo, além de verificar se o processo se encontra ou não fora de controle, é também necessário bom senso ao analisar qual a estimativa pontual (valor de referência) para a proporção de itens defeituosos. Pois, um processo sob controle que apresente uma alta proporção de itens defeituosos pode gerar muitos custos e prejuízos para a empresa, sendo muitas vezes um problema tão sério quanto a falta de controle do processo.

3.1.2 Gráfico p para amostras de tamanhos variáveis

Devido a fatores operacionais ou análises de custos referentes à coleta de dados, em determinadas situações faz-se necessário utilizar tamanhos variáveis de amostra.

Suponha, por exemplo, que sejam produzidas tampas de canetas de diferentes cores. Quando a quantidade de tampas de uma determinada cor suficiente para atender à demanda é produzida, passa-se então a produzir as tampas de outra cor. Dessa forma, as amostras são estabelecidas de acordo com a cor dos itens produzidos e o monitoramento da proporção de tampas com um formato inadequado pode ser feito em um contexto em que são coletadas amostras de tamanhos variáveis.

Nesta seção são abordados os usos do Gráfico p em amostras de tamanhos variáveis. Os dois tópicos a seguir referem-se à construção do Gráfico p com especificação (seção 3.1.2.1) e do Gráfico p sem especificação (seção 3.1.2.2.).

3.1.2.1 Gráfico p com especificação (amostras de tamanhos variáveis)

A construção do gráfico temporal, que contém os valores de \hat{p}_i na ordenada e os indicadores das amostras na abscissa, é feita de maneira análoga à que foi apresentada nas seções (3.1.1.1) e (3.1.1.2). A única diferença está no cálculo de \hat{p}_i , pois o tamanho das amostras varia.

Assim, o cálculo dos limites de controle para cada amostra, é dado por:

$$p \pm 3 \sqrt{\frac{p(1-p)}{n_i}}, \quad (3.1.2.1.a)$$

em que p é o valor de referência (especificado) da proporção de itens defeituosos.

Como o desvio-padrão de \hat{p}_i , neste caso, depende do tamanho das amostras, o Gráfico p para amostras de tamanhos variáveis pode não ser tão adequado. A grande dificuldade associada ao gráfico com limites de controle que variam de uma amostra para outra é que, conforme discussão realizada por Montgomery (2004, P.188 [12]), os testes para identificar “sequências ou padrões aparentemente anormais nos gráficos” não são indicados. Dessa forma, uma alta proporção amostral de itens defeituosos pode não ser tão significativa quanto uma proporção amostral um pouco menor calculada a partir de mais unidades amostrais. Uma saída para verificar a existência de padrões de repetição, consiste em utilizar uma variável padronizada z_i , dada por:

$$z_i = \frac{\hat{p}_i - p}{\sqrt{\frac{p(1-p)}{n_i}}} . \quad (3.1.2.1.b)$$

A variável z_i permite que se observe qual o desvio entre a proporção amostral de itens defeituosos e o valor de referência para essa proporção, em termos de desvios-padrão dessa estatística. Portanto, todos os valores z_i se encontram na mesma escala de medida, o que justifica que possam ser aplicados testes para verificar se existem sequências ou padrões de repetição (sazonalidades) no comportamento dessa variável aleatória.

Caso o processo se encontre sob controle, o valor esperado para a estatística z_i é igual a 0 e o seu desvio-padrão é igual a 1. Os limites de controle do gráfico da variável z_i (também conhecido como Gráfico p padronizado) são, portanto, iguais a ± 3 e a linha central é igual a 0.

3.1.2.2 Gráfico p sem especificação (amostras de tamanhos variáveis)

Nas situações em que não é possível obter o valor de referência de p através da empresa que deseja controlar a qualidade, é preciso estimá-lo considerando os tamanhos variáveis de amostra, através de:

$$\bar{p} = \frac{\sum_{i=1}^m D_i}{\sum_{i=1}^m n_i} , \quad (3.1.2.2.a)$$

em que D_i é o número de itens defeituosos na amostra i .

Os limites de controle para cada amostra, são então dados por:

$$\bar{p} \pm 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}} . \quad (3.1.2.2.b)$$

Caso haja interesse em verificar a existência de sequências ou padrões de repetição, deve-se construir o gráfico temporal da variável padronizada z_i , dada por:

$$z_i = \frac{\hat{p}_i - \bar{p}}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}} , \quad (3.1.2.2.c)$$

cujo valor esperado é igual a 0 e o desvio-padrão é igual a 1, caso o processo esteja sob controle. Portanto, a linha central do gráfico da variável padronizada z_i (ou Gráfico p padronizado) é igual a 0 e os limites de controle são iguais a ± 3 .

Um algoritmo para construção e análise do Gráfico p , considerando amostras de tamanhos iguais ou variáveis, é apresentado no Quadro 3.1.a do Apêndice C.

3.2 Gráfico np

O Gráfico np tem o objetivo de monitorar o número de itens defeituosos em cada amostra produzida em um determinado período de tempo. A grande vantagem desse gráfico é que, muitas vezes, ele é mais facilmente interpretado por pessoas que possuem pouco treinamento em Estatística, o que pode ser o caso, por exemplo, dos profissionais da linha de produção. Observe, contudo, que a fundamentação dos gráficos que monitoram a proporção amostral de defeitos ou o número de itens com defeitos é equivalente. No caso do Gráfico np , a estatística apresentada, isto é, aquela cujos valores estão na ordenada, é dada por:

$$n\hat{p}_i = \sum_{j=1}^n x_{ij} = D_i, \quad (3.2.a)$$

que representa o número de itens defeituosos na amostra i .

O valor esperado para essa estatística é dado por:

$$E(n\hat{p}_i) = E\left(\sum_{j=1}^n x_{ij}\right) = \sum_{j=1}^n E(x_{ij}) = np. \quad (3.2.b)$$

Similarmente, o desvio-padrão dessa estatística, devido à suposição de independência entre as observações presentes na mesma amostra, é dado por:

$$DP(n\hat{p}_i) = \sqrt{\text{Var}(n\hat{p}_i)} = \sqrt{\text{Var}\left(\sum_{j=1}^n x_{ij}\right)} = \sqrt{\sum_{j=1}^n \text{Var}(x_{ij})} = \sqrt{np(1-p)}. \quad (3.2.c)$$

Contudo, observe que tanto o valor esperado quanto o desvio-padrão de $n\hat{p}_i$ são funções de n . Dessa forma, caso fosse construído um gráfico em que cada amostra tivesse um tamanho diferente, a linha central e os limites de controle do gráfico iriam variar de uma amostra para outra, tornando, assim, complexa a sua interpretação. Por esta razão é que as metodologias de análise do Gráfico p são mais indicadas quando o tamanho das amostras coletadas é variável.

A seguir estão separados por, com especificação (seção 3.2.1) e sem especificação (seção 3.2.2), o detalhamento da construção desse gráfico.

3.2.1 Gráfico np com especificação

Quando está disponível uma especificação (valor de referência) para a proporção de defeitos de cada amostra, tem-se que a linha central do gráfico, de acordo com as equações (3.2.b) e (3.2.c), é dada por np e os limites de controle são dados por:

$$np \pm 3\sqrt{np(1-p)}. \quad (3.2.1.a)$$

3.2.2 Gráfico np sem especificação

Quando não é dada uma especificação (valor de referência) para a proporção de itens defeituosos em cada amostra, essa proporção é estimada a partir da equação (3.1.1.2.a), de tal forma que a linha central do gráfico seja dada por $n\bar{p}$ e os limites de controle sejam dados por:

$$n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}. \quad (3.2.2.a)$$

Um algoritmo para construção e análise do Gráfico np , considerando amostras de mesmo tamanho, é apresentado no Quadro 3.2.a do Apêndice C.

3.3 Exemplo 5

O primeiro conjunto de dados a ser analisado neste capítulo está exposto na Tabela 3.3 do Apêndice B e é constituído por 25 amostras de aparelhos celulares fabricados por uma empresa de telecomunicações qualquer (fictícia). Essas amostras representam o total, dentre 100 observações, de aparelhos com defeito (vício) de fabricação. Suponha que o interesse da empresa seja manter o processo sob controle, considerando uma especificação de 0,1 para a proporção de aparelhos defeituosos. A análise desse conjunto de dados é realizada, a seguir, através do Gráfico p .

Nas Figuras 3.3.a, 3.3.b, 3.3.c e 3.3.d é apresentada uma sequência de passos para a análise do conjunto de dados em questão no sistema de CEP *on-line* desenvolvido, desde a inserção dos dados (Figura 3.3.a) até a visualização do relatório instantâneo para o Gráfico *p* (Figura 3.3.d).

O Gráfico *p* apresentado na Figura 3.3.d detectou a última amostra como fora de controle, pois a proporção de aparelhos defeituosos nesta amostra é maior do que o limite superior de controle do gráfico. Outro fato importante é que as 6 últimas amostras estão acima da linha central do gráfico. Assim, o próximo passo seria tentar identificar quais são as causas responsáveis pela falta de controle do processo, visando garantir que a qualidade seja preservada no futuro. Para o caso em que não se considera a especificação do parâmetro, o comportamento do gráfico construído (que não é exibido aqui) é equivalente ao observado naquele presente na Figura 3.3.d.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Num defeituosos
14
7
11
13
6
5
10
13
9
9
11
12
```

Figura 3.3.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo5
Número de variáveis:	1
Número de observações:	2500
Número de amostras:	25
Tamanho das amostras:	100

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 3.3.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Proporção:

Avançar

Figura 3.3.c: Definindo as características e condições do processo.

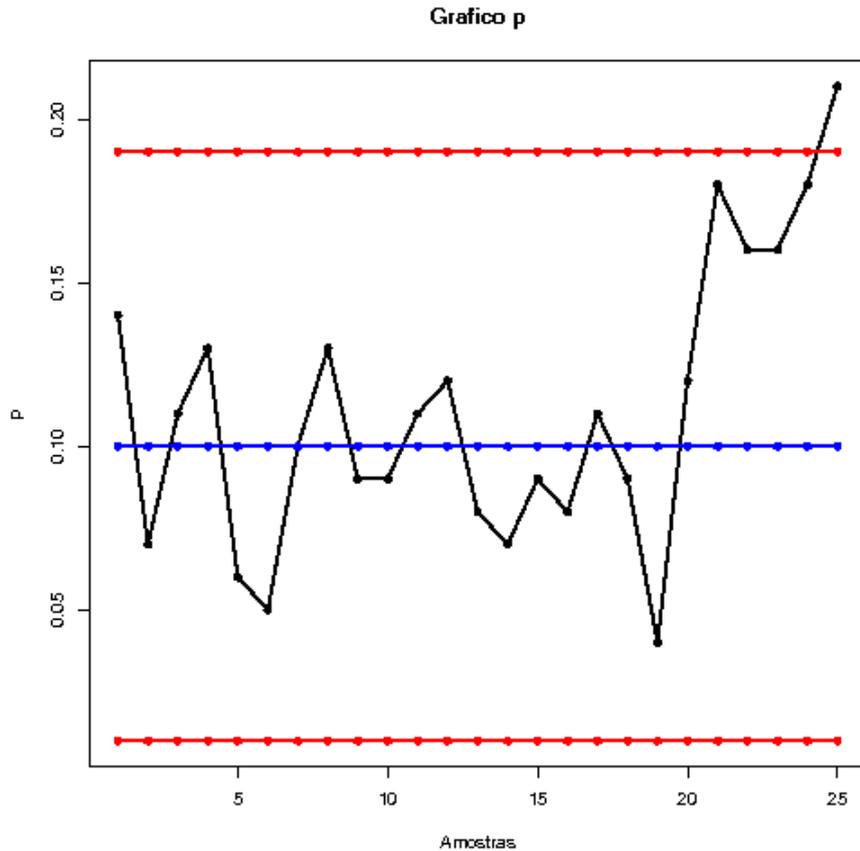


Figura 3.3.d: Relatório instantâneo gerado para o Gráfico p .

3.4 Exemplo 6

O segundo conjunto de dados analisado está exposto na Tabela 3.4 do Apêndice B e apresenta o total de aparelhos celulares defeituosos em cada uma das 25 amostras coletadas, bem como o total de aparelhos observados em cada uma dessas amostras. Neste exemplo, os tamanhos amostrais variam entre 90 e 140. Suponha que o interesse da empresa seja manter o processo sob controle, mas esta não tem especificações para a probabilidade p de que um aparelho qualquer produzido seja defeituoso. A análise é realizada, a seguir, através dos Gráficos p e p padronizado, visto que as amostras são de tamanhos variáveis.

Nas Figuras 3.4.a, 3.4.b, 3.4.c e 3.4.d são apresentados os passos necessários para a análise do conjunto de dados em questão no sistema de CEP *on-line* proposto, desde a inserção dos dados (Figura 3.4.a) até a visualização do relatório instantâneo para os Gráficos p e p padronizado (Figura 3.4.d).

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

```
101 94 108 127 113 137 136
90 137 93 130 130 125 129
140 96 98 140 104 137 139
111 111 137 119
```

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Num_defeituosos
12
11
10
7
14
11
14
6
17
10
15
14
```

Figura 3.4.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo6
Número de variáveis:	1
Número de observações:	2982
Número de amostras:	25
Tamanho das amostras:	vários

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 3.4.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Avançar

Figura 3.4.c: Definindo as características e condições do processo.

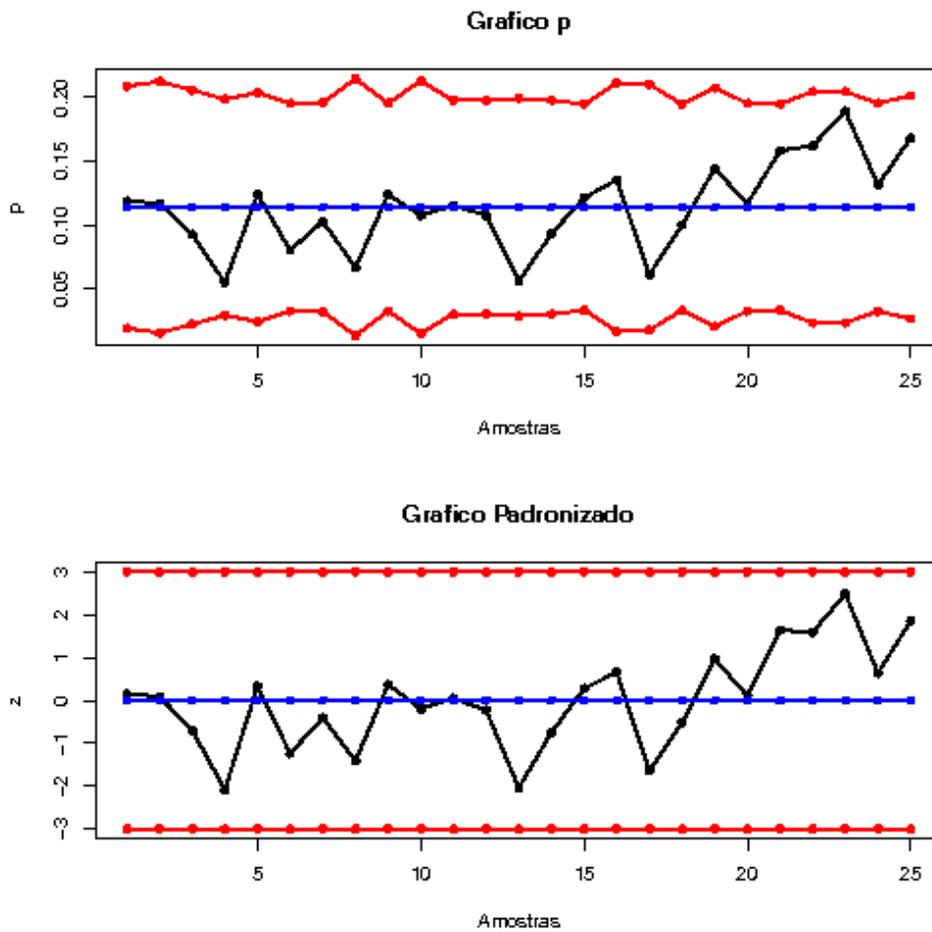


Figura 3.4.d: Relatório instantâneo gerado para os Gráficos p e p padronizado.

Como pode ser visto na Figura 3.4.d, não há pontos fora dos limites de controle em nenhum dos dois gráficos de controle da proporção de itens defeituosos. No entanto, observe que as 5 últimas amostras se encontram acima da linha central do gráfico e que as demais amostras estão localizadas, em sua maioria, abaixo dessa linha. Uma possível explicação para isto está no fato de que as 5 últimas amostras possuem os maiores valores observados para a proporção amostral de itens defeituosos e, por isso, acabam deslocando a linha central para cima (lembre-se que o estimador de p , neste caso, é uma média amostral, que é sensível a valores discrepantes). A análise (que não é realizada aqui) prossegue com a retirada das 5 últimas amostras do banco de dados e a reconstrução dos gráficos de controle a partir das demais, para saber se há mais amostras fora de controle.

3.5 Exemplo 7

Considere novamente o conjunto de dados analisado no Exemplo 5. No presente exemplo, a análise deste conjunto é realizada através do Gráfico np , utilizando ainda a especificação de 0,1 para a proporção de aparelhos defeituosos.

As Figuras 3.5.a, 3.5.b, 3.5.c e 3.5.d apresentam os passos necessários para a construção do Gráfico np para esse conjunto de dados no sistema de CEP *on-line*.

O Gráfico np apresentado na Figura 3.5.d é equivalente ao Gráfico p da Figura 3.3.d . Ou seja, as 5 últimas amostras se apresentam novamente como candidatas a fora de controle, por estarem todas bem acima da linha central do gráfico, inclusive com a última amostra excedendo o limite superior de controle. Para o Gráfico np construído sem considerar a especificação (este gráfico não é exibido aqui), o comportamento é equivalente ao observado naquele presente na Figura 3.5.d. Portanto, o próximo passo da análise consiste na identificação das causas responsáveis pela falta de controle do processo. Se tais causas forem identificadas, deve-se ainda retirar as 5 últimas amostras do banco de dados e reconstruir o Gráfico np a partir das observações restantes, para saber se há mais amostras fora de controle.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```

Num_defeituosos
14
7
11
13
6
5
10
13
9
9
11
12

```

Figura 3.5.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo7
Número de variáveis:	1
Número de observações:	2500
Número de amostras:	25
Tamanho das amostras:	100

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 3.5.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Proporção:

Figura 3.5.c: Definindo as características e condições do processo.

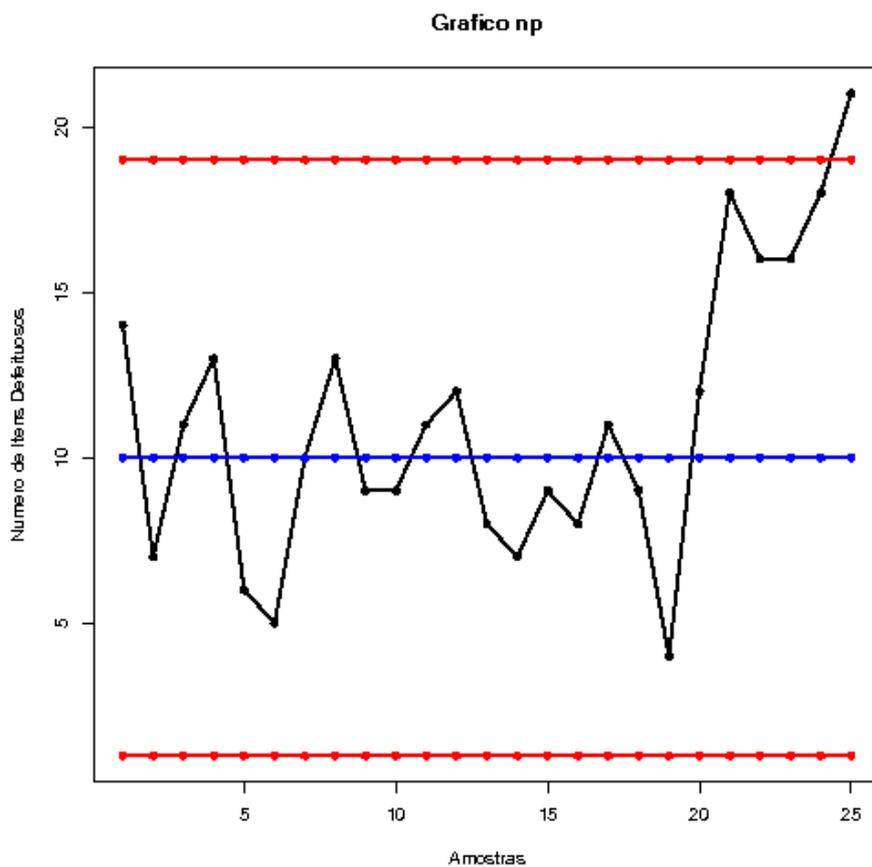


Figura 3.5.d: Relatório instantâneo gerado para o Gráfico *np*.

3.6 Considerações finais

Discutiu-se ao longo deste capítulo os gráficos de controle da proporção ou da frequência de itens defeituosos produzidos, isto é, os Gráficos p e np , que são utilizados quando se obtém variáveis qualitativas e deseja-se trabalhar com proporções de itens defeituosos, p . No entanto, em determinados casos, ao se produzir um equipamento, não é possível detectar defeitos em peças internas, pois esses defeitos não causam a falha total do produto. Por exemplo, na montagem de um computador em uma indústria ou comércio, é possível que uma memória de alto nível esteja funcionando com apenas metade de sua capacidade, porém este defeito não é detectado durante uma inspeção, uma vez que não prejudica o bom funcionamento do aparelho. Neste caso, os Gráficos p e np não podem ser usados. Uma saída para esse problema é a análise da qualidade através do controle do número de defeitos por unidade de inspeção. No próximo capítulo são apresentados dois gráficos, o Gráfico c e o Gráfico u , que podem ser utilizados nessas situações.

Capítulo 4

Metodologia CEP para atributos – Gráficos de controle do número de defeitos por item produzido

Embora um produto possa apresentar diversos defeitos, seu projeto pode garantir que ele mantenha um funcionamento adequado, como é o caso, por exemplo, de aparelhos eletrônicos. Muito embora diversos de seus componentes apresentem defeitos, a redundância presente no sistema é capaz de contornar tais defeitos e garantir o bom funcionamento do aparelho. Sendo assim, os Gráficos p e np não seriam capazes de identificar as variações no nível de qualidade do processo quando tais aparelhos fossem inspecionados. Nestes casos são necessários gráficos de controle para atributos úteis no monitoramento da frequência de defeitos por item. Entre tais gráficos, encontram-se o Gráfico c e o Gráfico u , os quais são discutidos em detalhes nas próximas seções.

4.1 Gráfico c

O Gráfico c é empregado no controle do número de defeitos por item produzido. É útil quando cada item produzido pode apresentar mais de um tipo de defeito ou repetições do mesmo defeito. A unidade inspecionada muitas vezes consiste de um item caro e tal inspeção demanda o investimento de tempo e dinheiro, fazendo com que cada uma dessas unidades (é possível também definir a unidade de inspeção como sendo um conjunto de n itens) seja considerada uma amostra. Naturalmente, todas as amostras devem ser constituídas de itens idênticos, para garantir que o número esperado de defeitos em uma determinada amostra não seja maior do que nas demais.

A construção do Gráfico c supõe que o número de defeitos por unidade inspecionada tenha distribuição Poisson com parâmetro c (c é a taxa de defeitos). Dessa forma, o número de defeitos possíveis, que pode ser observado em cada item, pode ser qualquer valor presente no conjunto infinito, dado por $\omega = \{0, 1, 2, \dots\}$. Além disso, a probabilidade de que um defeito seja observado em um determinado ponto qualquer do produto, deve ser constante para todos os possíveis pontos desse produto. Montgomery (2004, P.193 [12]) discute que, em contextos reais, essa suposição não costuma ser plenamente atendida, mas que ainda assim é possível que o modelo Poisson represente

uma boa aproximação do modelo responsável pelo número de defeitos por unidade inspecionada.

Seja X definida como o número de defeitos por unidade inspecionada. A suposição de que essa variável tenha distribuição Poisson com parâmetro c implica em que seu valor esperado seja igual a c e seu desvio-padrão seja igual a \sqrt{c} . É comum a política de qualidade da empresa, as exigências técnicas referentes à unidade produzida e as necessidades do cliente determinarem qual deve ser o valor esperado para o número de defeitos por unidade produzida. Nesse caso, o valor de referência (linha central) do gráfico é igual à própria especificação, c , e os limites de controle são iguais a:

$$c \pm 3\sqrt{c} . \quad (4.1.a)$$

Caso o limite inferior apresentado na equação (4.1.a) seja negativo, deve-se assumir este valor como nulo.

Por outro lado, se não for dada a especificação do processo, é necessário estimar o valor do parâmetro c , dado pelo número médio de defeitos observados,

$$\bar{c} = \frac{1}{m} \sum_{i=1}^m x_i , \quad (4.1.b)$$

em que m é o número de amostras e x_i é o total de defeitos na i -ésima amostra. Note que

\bar{c} é um estimador não-viesado de c , pois $E(\bar{c}) = E\left(\frac{1}{m} \sum_{i=1}^m x_i\right) = \frac{1}{m} \sum_{i=1}^m E(x_i) = E(x_i) = c$.

Dessa forma, os limites de controle do Gráfico c sem especificação, são dados por:

$$\bar{c} \pm 3\sqrt{\bar{c}} . \quad (4.1.c)$$

Um algoritmo para construção e análise do Gráfico c , considerando amostras de mesmo tamanho, é apresentado no Quadro 4.1.a do Apêndice C.

4.2 Gráfico u

O Gráfico u é utilizado quando se deseja monitorar o número médio de defeitos por unidade produzida. Uma de suas principais aplicações ocorre quando o planejamento amostral é alterado de tal forma que o tamanho das amostras passa então a ser multiplicado por uma constante κ , ou seja, uma amostra que inicialmente era de tamanho n , passa então a ser substituída por uma amostra de tamanho $n\kappa$. Observe que n e κ devem ser valores positivos e κ não necessariamente precisa ser inteiro.

Para construir os limites de controle do Gráfico u , faz-se necessário definir primeiramente algumas variáveis. Seja x_i o número de defeitos observados na i -ésima amostra após a alteração no planejamento amostral, m o número de amostras e n o tamanho de cada amostra. Suponha que x_i siga uma distribuição Poisson com parâmetro μ . Então, $E(x_i) = \mu = Var(x_i)$ e, sabendo que $\hat{u}_i = x_i / n$ é o número médio de defeitos na amostra i , tem-se que:

$$E(\hat{u}_i) = u = \mu / n. \quad (4.2.a)$$

Note que \hat{u}_i segue distribuição Poisson com parâmetro $u = \mu / n$. Portanto, um estimador não-viesado para u é dado por:

$$\tilde{u} = \bar{u} = \frac{1}{m} \sum_{i=1}^m \hat{u}_i, \quad (4.2.b)$$

pois $E(\bar{u}) = E\left(\frac{1}{m} \sum_{i=1}^m \hat{u}_i\right) = \frac{1}{m} \sum_{i=1}^m E\left(\frac{x_i}{n}\right) = \frac{1}{mn} \sum_{i=1}^m E(x_i) = \frac{E(x_i)}{n} = \frac{\mu}{n} = u.$

O estimador do desvio-padrão da variável \hat{u}_i , é dado por:

$$DP(\hat{u}_i) = \sqrt{u / n}. \quad (4.2.c)$$

Isso ocorre pois, caso o processo esteja sob controle, e utilizando a suposição de independência entre as n unidades pertencentes à mesma amostra, tem-se que

$$DP(\hat{u}_i) = \sqrt{Var(\hat{u}_i)} = \sqrt{Var\left(\frac{x_i}{n}\right)} = \sqrt{\frac{1}{n^2} Var(x_i)} = \sqrt{\frac{\mu}{n^2}}. \text{ Contudo, de acordo com a}$$

equação (4.2.a), $\mu = un$. Logo, $DP(\hat{u}_i) = \sqrt{\frac{un}{n^2}} = \sqrt{\frac{u}{n}}$. Caso u seja desconhecido, o

desvio-padrão de \hat{u}_i pode ser estimado através de:

$$\sqrt{\hat{V}ar(\hat{u}_i)} = \sqrt{\bar{u} / n}. \quad (4.2.d)$$

A seguir são abordadas duas situações do gráfico estudado: Gráfico u para amostras de mesmo tamanho (seção 4.2.2.1) e Gráfico u para amostras de tamanhos variáveis (seção 4.2.2.2).

4.2.1 Gráfico u para amostras de mesmo tamanho

Nesta seção é discutida a construção do Gráfico u para amostras de mesmo tamanho, quando é dada uma especificação para o parâmetro da distribuição da variável observada (seção 4.2.1.1), e também quando tal especificação não é dada (seção 4.2.1.2).

4.2.1.1 Gráfico u com especificação (amostras de mesmo tamanho)

Quando o local no qual está sendo avaliada a qualidade de um processo fornece uma especificação para o parâmetro, a construção do gráfico em questão segue o mesmo padrão das demais vistas até então. Assim, de acordo com as equações (4.2.a) e (4.2.c), a linha de referência (linha central) do gráfico deve ser dada pela especificação, isto é, por u e os limites de controle são dados por:

$$u \pm 3\sqrt{\frac{u}{n}}, \quad (4.2.1.1.a)$$

em que $\sqrt{\frac{u}{n}}$ é o desvio-padrão de \hat{u}_i , definido em (4.2.c).

4.2.1.2 Gráfico u sem especificação (amostras de mesmo tamanho)

Quando não é dada uma especificação que determine o valor ideal de u (ou para o número ideal de defeitos por unidade produzida), é preciso estimá-la. Assim, de acordo com as equações (4.2.b) e (4.2.d), a linha central do gráfico deve ser dada por \bar{u} (estimador do valor esperado para o número médio de defeitos por unidade produzida) e os limites de controle, por:

$$\bar{u} \pm 3\sqrt{\frac{\bar{u}}{n}}, \quad (4.2.1.2.a)$$

4.2.2 Gráfico u para amostras de tamanhos variáveis

O Gráfico c não é indicado para a realização do controle do número de defeitos por amostra quando estas diferem de tamanho. A razão para a não indicação é que a sua interpretação é demasiadamente complexa, pois a linha central do gráfico varia de uma amostra para outra. O Gráfico u é mais indicado nestes casos, uma vez que monitora o número médio de defeitos por unidade da amostra. Além disso, o Gráfico u possui uma linha central constante para todas as amostras, embora os limites de controle variem de acordo com o tamanho da amostra.

Nesta seção é discutida a construção do Gráfico u para amostras de tamanhos variáveis, quando se dispõe de uma especificação para o parâmetro da variável observada (seção 4.2.2.1) e também quando nenhuma especificação está disponível (seção 4.2.2.2).

4.2.2.1 Gráfico u com especificação (amostras de tamanhos variáveis)

Este caso é equivalente ao apresentado na seção 4.2.1.1, exceto por se diferenciar no cálculo do número médio de defeitos por amostra, que é $\hat{u}_i = x_i / n_i$.

Os limites de controle são dados por:

$$u \pm 3\sqrt{\frac{u}{n_i}}. \quad (4.2.2.1.a)$$

O Gráfico u , cujos limites de controle são apresentados na equação (4.2.2.1.a), assim como o Gráfico p para amostras de tamanhos variáveis, possui a vantagem de ser de fácil construção e compreensão. Contudo, os testes para verificação da presença de sequências e padrões de repetição no comportamento da variável aleatória \hat{u}_i não são indicados, visto que as escalas de variabilidade de \hat{u}_i e \hat{u}_j diferem sempre que $n_i \neq n_j$. Dessa forma, caso haja interesse na realização de tais testes, o gráfico da variável padronizada z_i (também conhecido como Gráfico u padronizado) é mais indicado. Essa variável é definida como:

$$z_i = \frac{\hat{u}_i - u}{\sqrt{u / n_i}}. \quad (4.2.2.1.b)$$

Caso o processo se encontre sob controle, a média e o desvio-padrão da variável z_i devem ser iguais a 0 e 1, respectivamente. Portanto, os limites de controle do gráfico da variável padronizada z_i são dados por ± 3 e a linha central, pelo valor 0.

4.2.2.2 Gráfico u sem especificação (amostras de tamanhos variáveis)

Um estimador não-viesado do valor esperado para o número médio de defeitos por unidade produzida, quando o tamanho das amostras varia, é dado por:

$$\bar{u} = \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m n_i}, \quad (4.2.2.2.a)$$

$$\text{pois } E(\bar{u}) = E\left(\frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m n_i}\right) = E\left(\frac{\sum_{i=1}^m n_i \hat{u}_i}{\sum_{i=1}^m n_i}\right) = \frac{\sum_{i=1}^m n_i E(\hat{u}_i)}{\sum_{i=1}^m n_i} = E(\hat{u}_i) = u.$$

Finalmente, os limites de controle são dados por:

$$\bar{u} \pm 3 \sqrt{\frac{\bar{u}}{n_i}}. \quad (4.2.2.2.b)$$

Caso haja interesse em verificar a existência de sequências e padrões de repetição (sazonalidades), deve-se construir o gráfico temporal da variável padronizada z_i , dada por:

$$z_i = \frac{\hat{u}_i - \bar{u}}{\sqrt{\bar{u} / n_i}}, \quad (4.2.2.2.c)$$

cuja média (ou valor esperado) é igual a 0 e o desvio-padrão é igual a 1. Portanto, a linha central do gráfico da variável padronizada z_i (também conhecido como Gráfico u padronizado) é igual a 0 e os limites de controle são iguais a ± 3 .

Um algoritmo para construção e análise do Gráfico u , considerando amostras de tamanhos iguais ou variáveis, é apresentado no Quadro 4.2.a do Apêndice C.

4.3 Exemplo 8

O primeiro conjunto de dados a ser analisado neste capítulo, está exposto na Tabela 4.3 do Apêndice B e é constituído por 25 amostras de 100 aparelhos celulares fabricados por certa empresa de telecomunicações (fictícia). Isso se deve ao fato de que tais aparelhos são comercializados em amostras com 100 unidades cada. A variável observada neste exemplo é o número de defeitos presentes em cada amostra (defeitos na carcaça e nas peças internas dos aparelhos, principalmente). O monitoramento da qualidade dos aparelhos produzidos pode ser realizado, neste caso, através do Gráfico *c*, pois pode haver interesse em verificar se a taxa com que ocorrem os defeitos nas amostras comercializadas é constante ao longo de um período de tempo, ou se sofre influência de determinadas variações do processo. Suponha ainda que o interesse da empresa seja manter o processo sob controle, mas esta não tem especificações em relação ao parâmetro.

Nas Figuras 4.3.a, 4.3.b, 4.3.c e 4.3.d é apresentada uma sequência de passos para a análise do conjunto de dados em questão no sistema de CEP *on-line*, desde a inserção dos dados (Figura 4.3.a) até a visualização do relatório instantâneo para o Gráfico *c* (Figura 4.3.d).

O Gráfico *c* apresentado na Figura 4.3.d identificou a penúltima amostra como fora de controle, pois o número total de defeitos nessa amostra é maior do que o limite superior de controle. Outro fato importante é que as 5 últimas amostras estão muito acima da linha central do gráfico. Assim, tais amostras acabam por influenciar no cálculo desta linha de referência, que é deslocada para cima, o que justificaria os valores muito baixos para o número total de defeitos observados na 5^a à 8^a amostra. A análise prossegue, buscando identificar as causas responsáveis pela falta de controle do processo e, em seguida, retiram-se as 5 últimas amostras do banco de dados e reconstrói-se o Gráfico *c* a partir das demais (tal gráfico não é exibido aqui), para saber se há mais amostras fora de controle.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```

Num_defeitos
13
10
10
12
5
7
2
6
12
6
14
9

```

Figura 4.3.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo8
Número de variáveis:	1
Número de observações:	2500
Número de amostras:	25
Tamanho das amostras:	100

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 4.3.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Figura 4.3.c: Definindo as características e condições do processo.

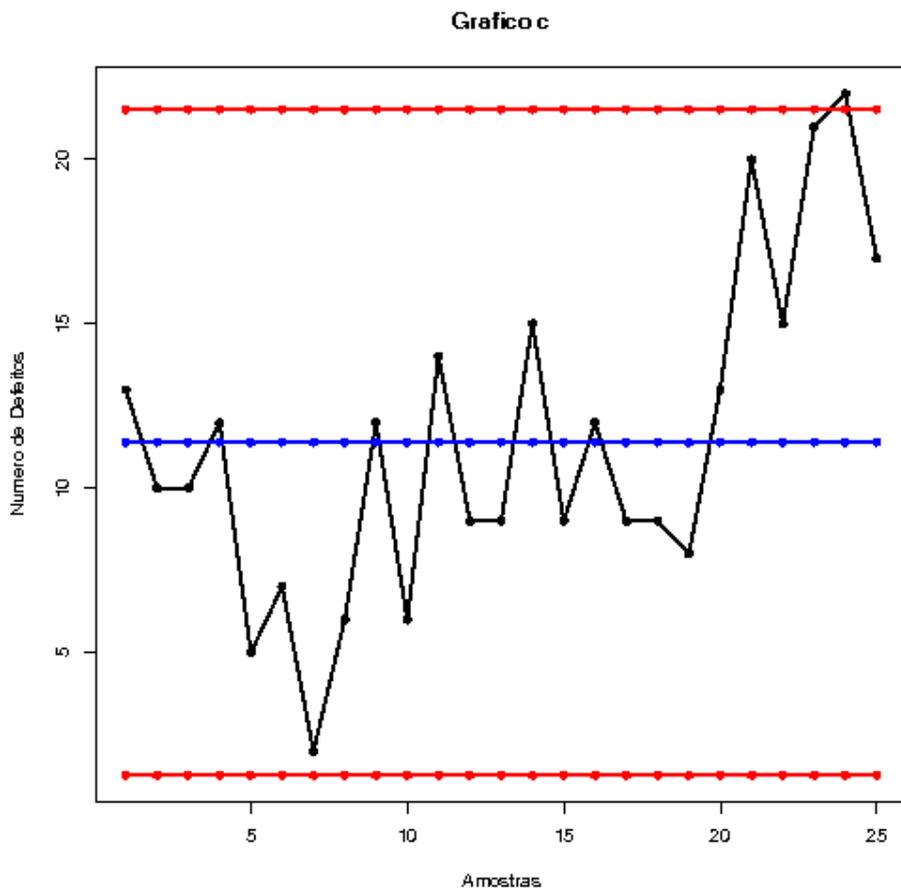


Figura 4.3.d: Relatório instantâneo gerado para o Gráfico c.

4.4 Exemplo 9

O segundo conjunto de dados analisado está exposto na Tabela 4.4 do Apêndice B. Neste conjunto é observado o número de defeitos (na carcaça, nas peças internas, etc.) nos aparelhos celulares fabricados pela empresa estudada nos exemplos anteriores. Visto que tais aparelhos são transportados em amostras de 100 unidades e que são reunidas 5 amostras destes aparelhos em uma caixa para a realização do transporte, há interesse em verificar se as caixas transportadas apresentam um padrão de qualidade aceitável. Uma forma de garantir a qualidade dos aparelhos apresentados neste conjunto de dados é através do controle do número médio de defeitos por amostra, realizado por meio do Gráfico u .

Nas Figuras 4.4.a, 4.4.b, 4.4.c e 4.4.d são apresentados os passos necessários para a análise do conjunto de dados em questão no sistema de CEP *on-line*, desde a inserção dos dados (Figura 4.4.a) até a visualização do relatório instantâneo para o Gráfico u (Figura 4.4.d).

O Gráfico u é apresentado na Figura 4.4.d para o caso em que é dada uma especificação para o valor esperado ideal de defeitos por amostra ($u=10$). Tal gráfico identificou a falta de controle das 5 últimas amostras, visto que tais amostras estão localizadas muito acima da linha de referência (ou linha central) do gráfico, sendo que duas delas (as amostras de números 22 e 23) excedem o limite superior de controle. Portanto, é importante investigar quais foram os fatores que conduziram ao resultado observado na Figura 4.4.d, visando identificar se existiram causas que possam ser controladas. O comportamento do Gráfico u para o caso em que não é considerada a especificação do parâmetro (tal gráfico não é mostrado aqui) é equivalente ao observado na Figura 4.4.d.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```

Num defeitos
14
8
12
16
15
6
9
11
10
7
6
7

```

Figura 4.4.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo9
Número de variáveis:	1
Número de observações:	125
Número de amostras:	25
Tamanho das amostras:	5

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 4.4.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Número Médio de Defeitos:

Figura 4.4.c: Definindo as características e condições do processo.

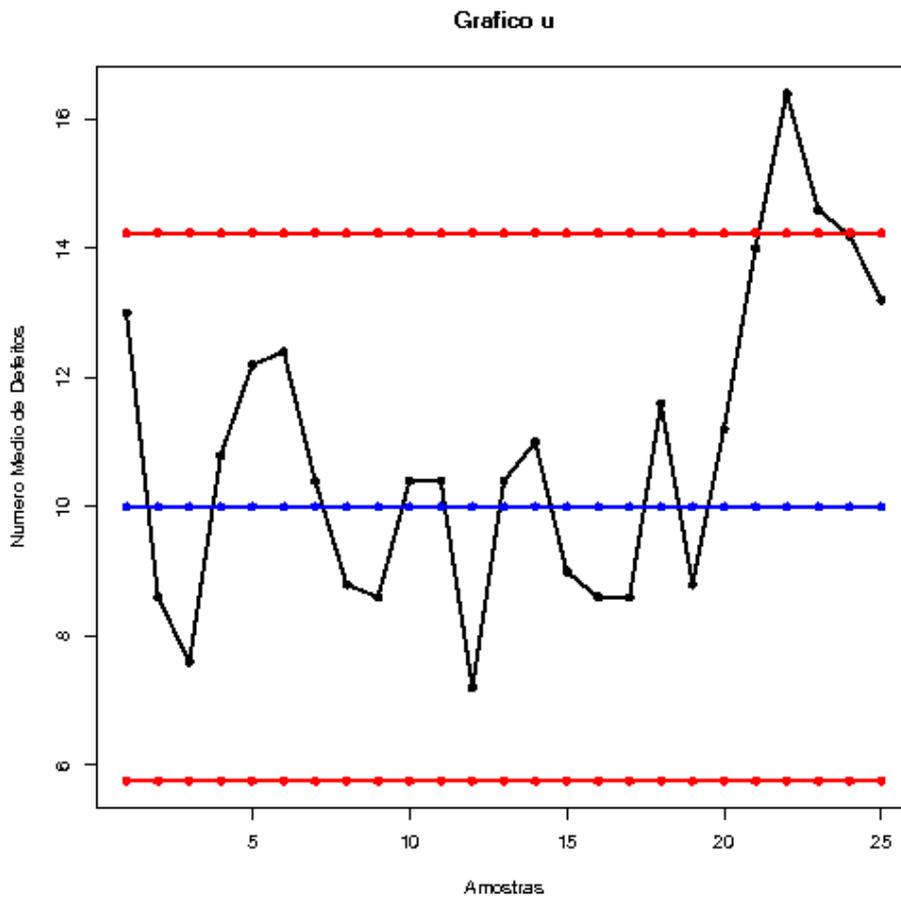


Figura 4.4.d: Relatório instantâneo gerado para o Gráfico *u*.

4.5 Exemplo 10

Suponha que, devido a uma alteração proposta pelo setor de logística da empresa estudada nos exemplos anteriores, o transporte dos aparelhos celulares não é mais realizado em caixas com 5 amostras cada. São, a partir de então, utilizadas caixas que contêm entre 4 e 7 amostras (cada amostra contém 100 aparelhos). O conjunto de dados que está exposto na Tabela 4.5 do Apêndice B é formado por amostras destas caixas. Dessa forma, observa-se o total de defeitos presentes em cada amostra formada por uma dessas caixas de tamanho variável.

Visando promover um padrão de qualidade constante nos aparelhos presentes nas caixas de transporte utilizadas pela empresa, duas alternativas úteis são o Gráfico u e o Gráfico u padronizado. Esta escolha se deve ao fato de que as caixas de transporte possuem tamanho variável.

Nas Figuras 4.5.a, 4.5.b, 4.5.c e 4.5.d é apresentada uma sequência de passos para a análise do conjunto de dados em questão no sistema de CEP *on-line* desenvolvido, desde a inserção dos dados (Figura 4.5.a) até a visualização do relatório instantâneo para os Gráficos u e u padronizado (Figura 4.5.d).

A Figura 4.5.d apresenta os Gráficos u e u padronizado para o caso em que é considerada a especificação para o número ideal de defeitos por amostra transportada ($u=10$). Ambos os gráficos identificaram as 5 últimas amostras como fora de controle, uma vez que seus números médios de defeitos são muito superiores à linha de referência (ou linha central) e 4 destes (referentes às quatro últimas amostras) excedem o limite superior de controle. Para o caso em que não é considerada a especificação estabelecida para o processo, o comportamento dos gráficos (que não são mostrados aqui) é equivalente ao observado naqueles apresentados na Figura 4.5.d. No entanto, a amostra de número 16 passa a ser considerada também como fora de controle, pois o número médio de defeitos nessa amostra é menor do que o limite inferior de controle dos gráficos construídos sem considerar a especificação.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

```
4 4 7 6 6 4 4 4 4 7 5 4 6 4
5 7 6 6 5 6 4 6 7 4 4
```

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Num defeitos
12
8
7
9
12
9
4
5
5
12
6
17
```

Figura 4.5.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo10
Número de variáveis:	1
Número de observações:	129
Número de amostras:	25
Tamanho das amostras:	vários

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 4.5.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu:

Escolha os gráficos adequados para seu tipo de dados:

Defina as condições para cada gráfico:

Número Médio de Defeitos:

Figura 4.5.c: Definindo as características e condições do processo.

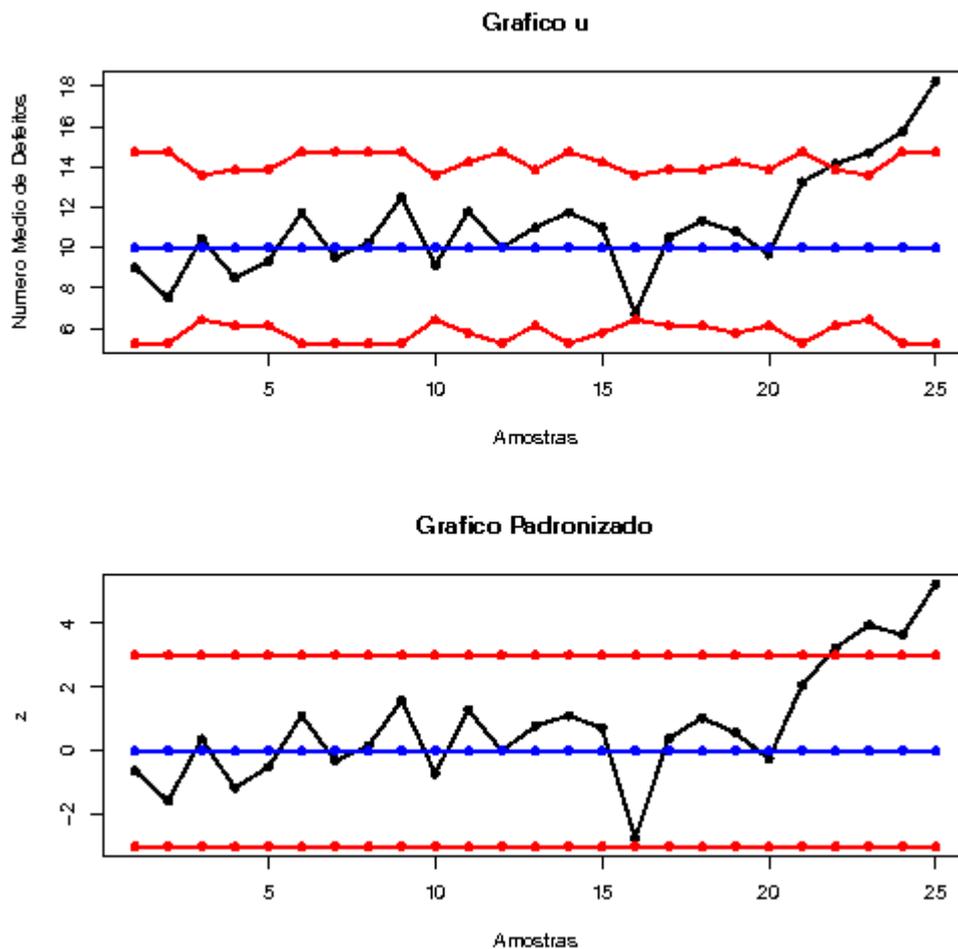


Figura 4.5.d: Relatório instantâneo gerado para os Gráficos u e u padronizado.

Visto que há indícios de que as 5 últimas amostras se encontram fora de controle, estas devem ser retiradas do banco de dados, para que os gráficos de controle possam ser reconstruídos (tais gráficos não são exibidos aqui) sem sua influência e seja possível verificar se a 16^a amostra se encontra, de fato, fora de controle. Tal escolha se deve ao fato de que há, de um lado, 5 amostras candidatas a fora de controle e, do outro, apenas 1. Dessa forma, é mais sugestivo que as 5 últimas amostras afetem o resultado da 16^a, em vez do contrário.

4.6 Considerações finais

Foram discutidos nos capítulos 3 e 4 os gráficos de controle para atributos, que são capazes de identificar se os valores esperados para a frequência ou probabilidade com que se produzem itens defeituosos e para a frequência com que se observam defeitos em uma determinada quantidade de itens, são constantes ao longo do tempo ou se sofrem influência de variações identificáveis do processo. É possível também comparar os parâmetros observados no processo com seus respectivos valores ideais através de gráficos construídos a partir de uma especificação. Tal comparação também foi feita no capítulo 2, em que se considerou o caso em que uma variável quantitativa é observada nos itens resultantes do processo.

Observou-se que os gráficos discutidos até agora foram capazes de fornecer relevantes informações referentes ao processo observado, conferindo ao analista condições de identificar quais foram os períodos em que ocorreram deslizes significativos no nível de qualidade. A partir destas informações, é possível promover futuramente um padrão de qualidade mais homogêneo ao longo do tempo, diminuindo assim custos e aumentando a satisfação do cliente.

Contudo, embora seja possível ter indicativos em relação à capacidade (ou capacidade) do processo de atender às especificações, necessidades e exigências do consumidor, os gráficos discutidos não são capazes de mensurar essa capacidade. Dessa forma, os índices de capacidade, que são discutidos no próximo capítulo, exercem um papel de grande relevância, pois conferem uma medida numérica à capacidade do processo de satisfazer tais necessidades e exigências do consumidor, no caso em que as variáveis observadas em cada unidade produzida são quantitativas.

Capítulo 5

Capacidade do processo

Nos capítulos anteriores foram discutidas algumas das principais ferramentas de prevenção e manutenção do controle estatístico de processos: os gráficos de controle para variáveis e atributos, que objetivam atender continuamente às condições ou necessidades exigidas pelo cliente ou aos limites de variação calculados, sempre com o intuito de monitorar o processo. Os gráficos de controle também podem indicar quão capaz é um processo para atender às especificações ou aos limites de variação calculados, embora existam técnicas mais objetivas para medir esta capacidade do processo. Dentre tais técnicas, destacam-se os índices de capacidade, devido à relevância de sua interpretação e à sua facilidade de utilização. Tais índices permitem que se estime, a partir de uma amostra aleatória representativa, qual é o nível de capacidade que o processo dispõe para atender às especificações estabelecidas para os itens produzidos.

Além da simplicidade dos cálculos necessários à sua construção, tais índices possuem a vantagem de resumir em um único valor a condição do processo, conferindo condições de se tomar rapidamente decisões embasadas em resultados estatísticos, muito embora também sejam ponderados durante a tomada de decisões os mais diversos fatores, como políticas de marketing e critérios financeiros.

Contudo, antes de se apresentar os principais índices de capacidade, é importante ressaltar que tais índices são construídos a partir de algumas suposições que necessariamente devem ser verificadas e respeitadas, a fim de garantir a confiabilidade da interpretação dos resultados observados.

Em primeiro lugar, os índices devem ser calculados a partir de amostras que estejam sob controle, isto é, amostras que tenham seus parâmetros constantes ao longo do tempo. Caso algumas amostras apresentem uma variância menor do que a usual durante certo período, a capacidade do processo pode ser superestimada, aumentando a devolução de itens fora de especificação e levando ao comprometimento da fidelidade e satisfação dos clientes. Por outro lado, se algumas amostras apresentarem variabilidade maior do que a verdadeira, pode-se concluir que os itens produzidos não são capazes de atender às especificações do cliente, gerando assim desnecessário aumento nos custos decorrentes das políticas de qualidade, na tentativa de produzir itens mais homogêneos.

Outra relevante suposição referente aos índices de capacidade é de que a variável observada segue uma distribuição Normal. Caso essa suposição não seja atendida, podem ocorrer erros substanciais na estimação dos índices, fazendo com que sua interpretação fique comprometida. Esses erros, de acordo com discussão realizada por Montgomery (2004, P.226 [12]), podem ocorrer mesmo nos casos em que a distribuição da variável é simétrica. Uma alternativa que pode contornar a violação dessa suposição e conduzir à utilização adequada dos índices de capacidade consiste em transformar os dados de forma a obter uma variável cuja distribuição possa ser bem aproximada pela Normal.

As duas suposições iniciais apresentadas implicam que, para se calcular a capacidade de um processo, primeiro é necessário realizar o CEP, com os gráficos de controle do processo, mesmo porque para a construção desses gráficos também é necessária a suposição de normalidade.

Finalmente, a terceira suposição de grande relevância é de que, no caso de especificações bilaterais, se a média do processo não corresponder ao centro dos limites de especificação, esta deve ao menos estar entre estes limites, para que os cálculos da capacidade sejam válidos. As Figuras 5.a e 5.b ilustram os dois possíveis casos nos quais essa suposição é violada. No gráfico apresentado na Figura 5.a, observa-se uma variável cuja média é inferior ao *LIE* (limite inferior de especificação) e no gráfico apresentado na Figura 5.b, uma variável cuja média é superior ao *LSE* (limite superior de especificação). Em ambos os casos, os índices de capacidade não devem ser utilizados.

Assim, listadas as suposições fundamentais para o bom desempenho dos índices a serem apresentados, resta discutir quais são suas principais aplicações, os métodos de estimação e a respectiva interpretação de cada um deles. Tudo isso é feito nas seções seguintes.

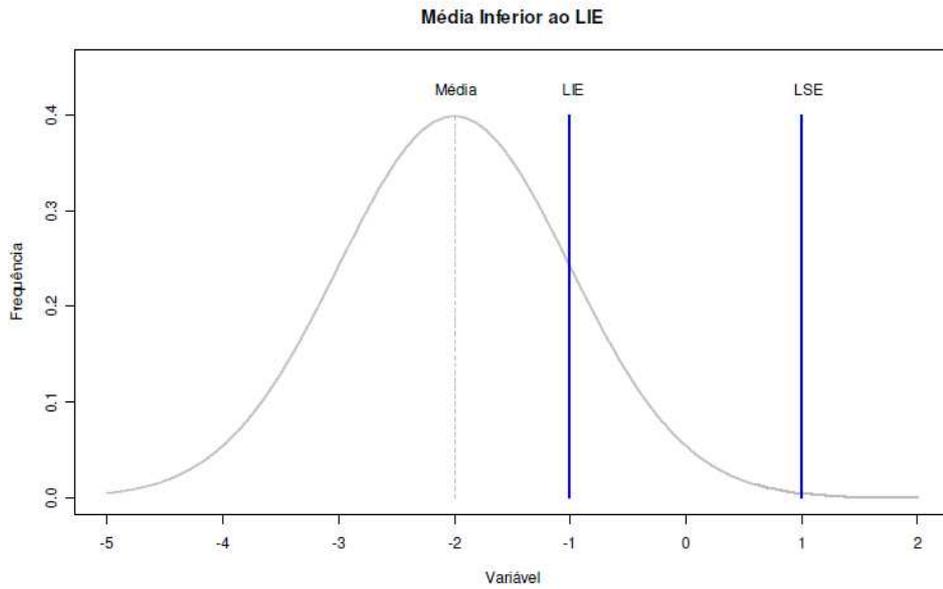


Figura 5.a: Processo com média abaixo do *LIE*.

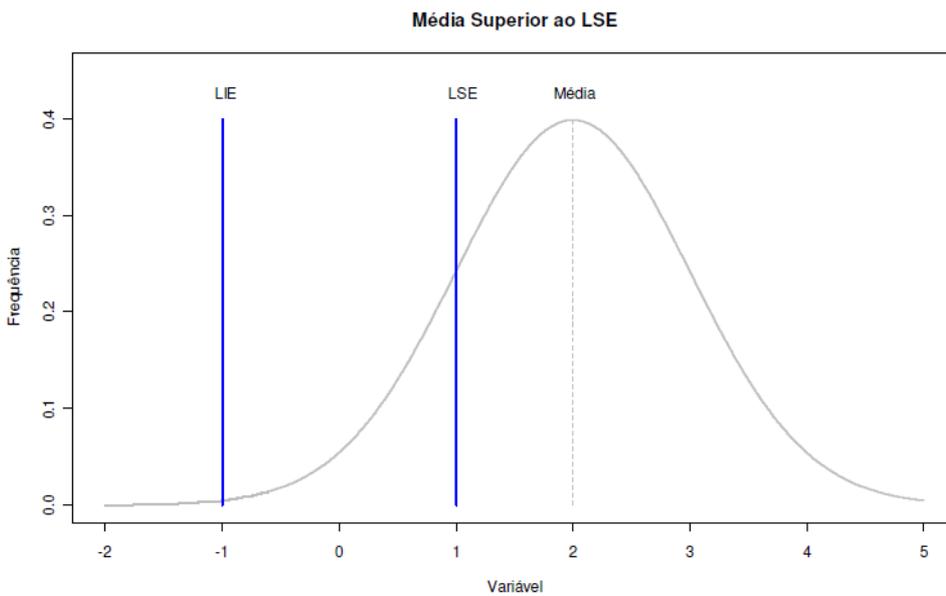


Figura 5.b: Processo com média acima do *LSE*.

5.1 Índices para um processo centrado

O índice de capacidade mais simples a ser apresentado neste capítulo consiste no C_p , que é utilizado quando é adequado supor que a média da variável mensurada coincide com o centro do intervalo de especificação (quando isso ocorre, diz-se que o

processo é centrado). A ideia deste índice é comparar a variabilidade natural do processo (igual a 6σ , em uma política de qualidade 6-sigma) com a variabilidade aceitável para o mesmo (igual à amplitude do intervalo de especificação, *LSE-LIE*). Este índice é, portanto, dado por:

$$C_p = \frac{LSE - LIE}{6\sigma}. \quad (5.1.a)$$

Quanto menor (inferior a 1) for o valor desse índice menor é a capacidade do processo, pois a variabilidade natural do processo deve ser menor que a variabilidade aceitável. Segundo Montgomery (2004, P.227 [12]), quando o valor de $C_p = 1$, significa que existe uma taxa de falhas de 27 partes por 10000, ou 2700 ppm (partes por milhão), considerando que os dados representam uma distribuição Normal e que os limites tenham 99,73% de confiança.

Contudo, muitas vezes é preciso estimar o valor de σ , de forma que o valor utilizado de fato na tomada de uma decisão, em geral, consiste de uma estimativa do valor de C_p . Esta estimativa é dada por:

$$\hat{C}_p = \frac{LSE - LIE}{6S}, \quad (5.1.b)$$

em que S representa o desvio-padrão amostral da variável observada X , dado por

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}.$$

Em função do fato de se dispor de uma estimativa para o valor do índice de capacidade do processo, a estimação por meio de um intervalo é um procedimento que permite fazer inferências tendo em vista um determinado nível de confiança. Em Montgomery (2004, P.231 [12]), o intervalo de confiança, ao nível de $100(\alpha)\%$, para o índice C_p é dado por:

$$\left(\hat{C}_p \sqrt{\frac{\chi_{n-1; (1-\alpha)/2}^2}{n-1}}; \hat{C}_p \sqrt{\frac{\chi_{n-1; (1+\alpha)/2}^2}{n-1}} \right) \quad (5.1.c)$$

(Kane, 1986 [9]), em que $\chi_{n-1; (1-\alpha)/2}^2$ e $\chi_{n-1; (1+\alpha)/2}^2$ representam, respectivamente, o $100(1-\alpha)/2$ -ésimo e o $100(1+\alpha)/2$ -ésimo quantil de uma distribuição Qui-quadrado com $n-1$ graus de liberdade.

Caso se assuma um valor de $\alpha = 0,95$, por exemplo, a interpretação do resultado da equação (5.1.c) é de que se fossem coletadas 100 diferentes amostras e para cada uma delas fosse construído o intervalo utilizando essa equação, aproximadamente 95 intervalos calculados iriam conter o verdadeiro valor de C_p . Portanto, tem-se que as chances de que esse intervalo contenha o verdadeiro C_p são de 95 em 100 (ou 95%).

A partir do índice C_p é possível ainda construir outro índice cuja interpretação é de grande relevância. Este índice é dado por:

$$P = 100 \left(\frac{1}{C_p} \right), \quad (5.1.d)$$

que é simplesmente a porcentagem da faixa de especificação (ou de tolerância) que é ocupada pelo processo. Assim, um valor de $C_p > 1$ significa que o processo ocupa menos de 100% da faixa de especificação. Consequentemente, relativamente poucos itens defeituosos são produzidos por esse processo. Se $C_p = 1$, então o processo ocupa toda a faixa de tolerância, o que implica em aproximadamente 0,27% (ou 2700 ppm) de itens defeituosos. Finalmente, $C_p < 1$ significa que o processo ocupa mais de 100% da faixa de tolerância. Neste caso, um grande número de itens defeituosos é produzido por esse processo.

Nas situações em que se dispõe de uma estimativa para C_p , obtém-se também uma estimativa para P , dada por:

$$\hat{P} = 100 \left(\frac{1}{\hat{C}_p} \right). \quad (5.1.e)$$

Assim, um intervalo de confiança, ao nível de $100(\alpha) \%$, para o índice P pode ser obtido diretamente a partir do intervalo para C_p , definido em (5.1.c), simplesmente tomando o inverso dos limitantes inferior e superior desse intervalo e, em seguida, multiplicando por 100.

5.2 Índices para um processo não-centrado

Nas situações em que não é plausível supor que a média do processo coincida com o centro do intervalo de especificação (quando isso ocorre, diz-se que o processo é não-centrado), torna-se necessário utilizar um índice de capacidade mais sofisticado do que o C_p . Uma alternativa útil é o índice dado por:

$$C_{pk} = \min\left(\frac{LSE - \mu}{3\sigma}; \frac{\mu - LIE}{3\sigma}\right), \quad (5.2.a)$$

em que μ é a média do processo e \min representa a função de mínimo.

A ideia para a construção desse índice consiste em considerar o menor desvio entre o valor central do processo, μ , e os limites de tolerância para o posicionamento da variável, comparando este valor obtido à variabilidade aceitável para o processo. Novamente, caso este índice assuma valores pequenos (inferiores a 1), tem-se que o processo dispõe de poucas condições para atender às especificações estabelecidas, sendo necessário promover a diminuição na variabilidade do processo ou a alteração em seu posicionamento central.

Contudo, nem sempre os valores de μ e σ são dados. Então, nesses casos, é

necessário usar as suas estimativas, dadas por $\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ e $\hat{\sigma} = S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$,

respectivamente. Assim, o cálculo da capacidade para processos não-centrados, a partir de valores estimados, é dado por:

$$\hat{C}_{pk} = \min\left(\frac{LSE - \bar{x}}{3S}; \frac{\bar{x} - LIE}{3S}\right). \quad (5.2.b)$$

Assim como para os processos centrados, pode-se calcular o intervalo de confiança para C_{pk} , que utiliza os valores $Z_{(1+\alpha)/2}$ (percentil de ordem $100(1+\alpha)/2$ % da distribuição Normal Padrão) e, segundo Montgomery (2004, P.231 [12]), é dado por:

$$\left(\hat{C}_{pk} \left[1 - Z_{(1+\alpha)/2} \sqrt{\frac{1}{9n\hat{C}_{pk}^2} + \frac{1}{2(n-1)}} \right]; \hat{C}_{pk} \left[1 + Z_{(1+\alpha)/2} \sqrt{\frac{1}{9n\hat{C}_{pk}^2} + \frac{1}{2(n-1)}} \right] \right). \quad (5.2.c)$$

Para maiores detalhes a respeito da construção desse intervalo, veja Kushler e Hurley (1992 [11]).

Contudo, embora a ideia para o desenvolvimento do índice C_{pk} esteja relacionada ao fato de que determinados processos apresentam uma média que difere do centro do intervalo de especificação, este índice não representa adequadamente a descentralização destes.

A Figura 5.2.a apresenta dois processos cujas médias, assim como também as variâncias, diferem entre si, porém cujos valores de C_{pk} são os mesmos. Um dos processos apresenta média igual a 1,5 e desvio-padrão igual a 1/3 e o outro apresenta média igual a 0,5 e desvio-padrão unitário. Os limites do intervalo de especificação são -2 e 2. Ambos os processos apresentam um C_{pk} igual a 0,5, embora tenham uma medida de posição central bem diferente.

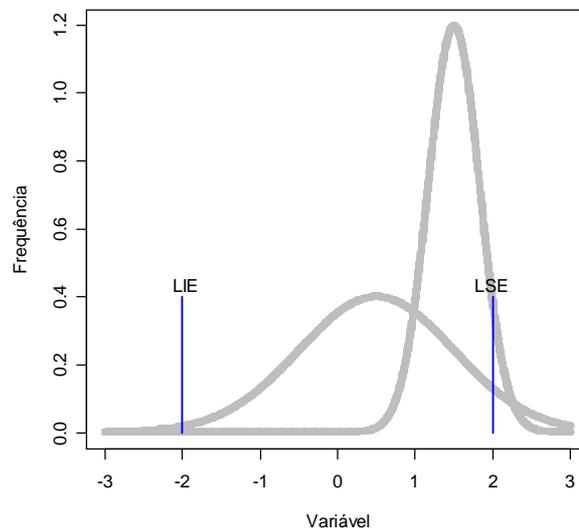


Figura 5.2.a: Dois processos distintos com mesmo C_{pk} .

Dessa forma, a utilização de um índice mais sofisticado, capaz de identificar melhor o posicionamento do processo, mostra-se conveniente caso se deseje uma compreensão clara das características do processo observado.

O índice C_{pm} constitui uma alternativa que atende ao objetivo de fornecer um quadro mais claro em relação ao posicionamento do processo. Esse índice é dado por:

$$C_{pm} = \frac{LSE - LIE}{6\eta}, \quad (5.2.d)$$

em que η é a raiz quadrada do erro quadrático médio da variável em estudo, X , em relação ao valor-alvo do processo, $T = \frac{1}{2}(LSE + LIE)$, isto é,

$$\eta^2 = E[(x-T)^2] = E[(x-\mu)^2] + (\mu-T)^2 = \sigma^2 + (\mu-T)^2. \quad (5.2.e)$$

Portanto, a equação (5.2.d) pode ser reescrita como:

$$C_{pm} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + (\mu-T)^2}} = \frac{C_p}{\sqrt{1+\xi^2}}, \quad (5.2.f)$$

em que $\xi = \frac{\mu-T}{\sigma}$.

Uma forma de estimar o C_{pm} , quando não são dados os valores de μ e σ , é através de:

$$\hat{C}_{pm} = \frac{\hat{C}_p}{\sqrt{1+V^2}}, \quad (5.2.g)$$

em que $V = \frac{\bar{x}-T}{S}$.

De acordo com estudos e discussão realizada por Boyles (1991 [3]), os valores dos índices C_{pk} e C_{pm} coincidem com o valor do C_p quando $\mu = T$ (isto é, quando o processo é centrado), e diminuem quando μ se afasta de T (isto é, quando o processo é não-centrado). Além disso, Boyles (1991 [3]) mostra também que o C_{pm} de um processo com $|\mu-T| = \Delta > 0$ é estritamente limitado acima pelo valor C_p de um processo com $\sigma = \Delta$, ou seja,

$$C_{pm} < \frac{LSE - LIE}{6|\mu-T|}. \quad (5.2.h)$$

A partir da equação (5.2.h), tem-se que uma condição necessária para que $C_{pm} \geq k$, k arbitrário, é $|\mu-T| < \frac{1}{6k}(LSE - LIE)$. Portanto, a partir do valor calculado para C_{pm} , é possível ter uma estimativa para a diferença máxima entre a média do processo, μ , e o valor ideal para essa média, T . Caso esta diferença máxima seja alta, pode ser necessário melhorar o posicionamento do processo a fim de atender às exigências estabelecidas.

5.3 Exemplo 11

O conjunto de dados analisado neste capítulo está exposto na Tabela 5.3 do Apêndice B e contém 25 amostras com 15 observações cada, representando a capacidade de armazenamento (em megabytes, MB) de cartões de memória para aparelhos celulares. Suponha que o interesse da empresa que produz esses cartões é manter o processo sob controle, mas esta não tem referências (especificações) para os parâmetros (média e desvio-padrão). Caso o processo esteja sob controle, há interesse também em avaliar o nível de capacidade que o mesmo dispõe para atender às exigências impostas pelo projeto do produto, cujos limites de especificação, que representam o intervalo de variação ideal para as capacidades de armazenamento observadas, foram dados entre 90 MB e 150 MB.

Nas Figuras 5.3.a, 5.3.b, 5.3.c, 5.3.d, 5.3.e e 5.3.f é apresentada uma sequência de passos para a análise desse conjunto de dados no sistema de CEP *on-line* desenvolvido, desde a inserção dos dados (Figura 5.3.a) até a visualização do relatório instantâneo para a capacidade (Figura 5.3.f).

Na Figura 5.3.d observa-se que o processo se encontra sob controle, tanto em relação ao gráfico da variabilidade (Gráfico *S*) quanto ao gráfico para a média (Gráfico \bar{X}), visto que não são observados, em nenhum dos dois gráficos, pontos fora dos limites de controle e nem tendências ou sazonalidades. Portanto, pode-se realizar o estudo da capacidade desse processo.

O histograma das observações sob controle, os índices de capacidade calculados e seus respectivos intervalos de 95% de confiança, são apresentados na Figura 5.3.f. Quanto ao histograma apresentado, observa-se que as capacidades dos cartões produzidos se distribuem em torno de 100 MB. De fato, a média amostral desse conjunto de dados é igual a 99,9 MB e o desvio-padrão amostral é igual a 9,8 MB. Como a média amostral difere do centro do intervalo de especificação (120 MB), o processo em questão é não-centrado. Logo, os índices C_{pk} e C_{pm} são os mais apropriados para a análise da capacidade. Tais índices assumem valores pequenos (inferiores a 1), que são indicativos de que o processo não é capaz de atender às especificações estabelecidas. Neste caso, deve-se não apenas diminuir a variabilidade do processo em questão, como também melhorar o posicionamento central do mesmo, uma vez que a

diferença máxima entre a média amostral e o valor ideal dessa média (que corresponde ao centro do intervalo de especificação) é alta (desvio = 22,37).

Se os limites de especificação da variável observada (capacidade de armazenamento) fossem dados entre 70 MB e 130 MB, os resultados da capacidade seriam aqueles apresentados na Figura 5.3.g. Note que, neste caso, a média amostral do conjunto de dados em questão (99,9 MB) coincide com o centro do intervalo de especificação (100 MB). Então, o processo é centrado (como pode ser visto também no histograma da Figura 5.3.g) e os índices C_p e P são os mais indicados para a análise. No caso, o índice C_p assume um valor próximo de 1, que é similar aos valores apresentados pelos índices C_{pk} e C_{pm} . Já o valor do índice P é próximo de 100 %, o que significa que o processo em questão ocupa praticamente toda a faixa de especificação (ou de tolerância). Portanto, é necessário que se promova uma diminuição na variabilidade desse processo, pois o mesmo encontra-se muito próximo de não ser capaz de atender às especificações estabelecidas.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (se n é constante):

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

Defina o nome da sua tabela de dados:

Cole aqui seus dados, sendo a primeira linha o nome da variável:

```
Peso
101.2
81.88
101.52
88.81
100.02
111.89
94.95
99.01
103.05
101.98
99.51
91.57
```

Avançar

Figura 5.3.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo11
Número de variáveis:	1
Número de observações:	375
Número de amostras:	25
Tamanho das amostras:	15

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 5.3.b: Confirmando os dados inseridos.

Escolha o tipo de dados que você inseriu: Variáveis

Escolha os gráficos adequados para seu tipo de dados: Gráficos S e Xbarra

Defina as condições para cada gráfico: Sem especificação

Avançar

Figura 5.3.c: Definindo as características e condições do processo.

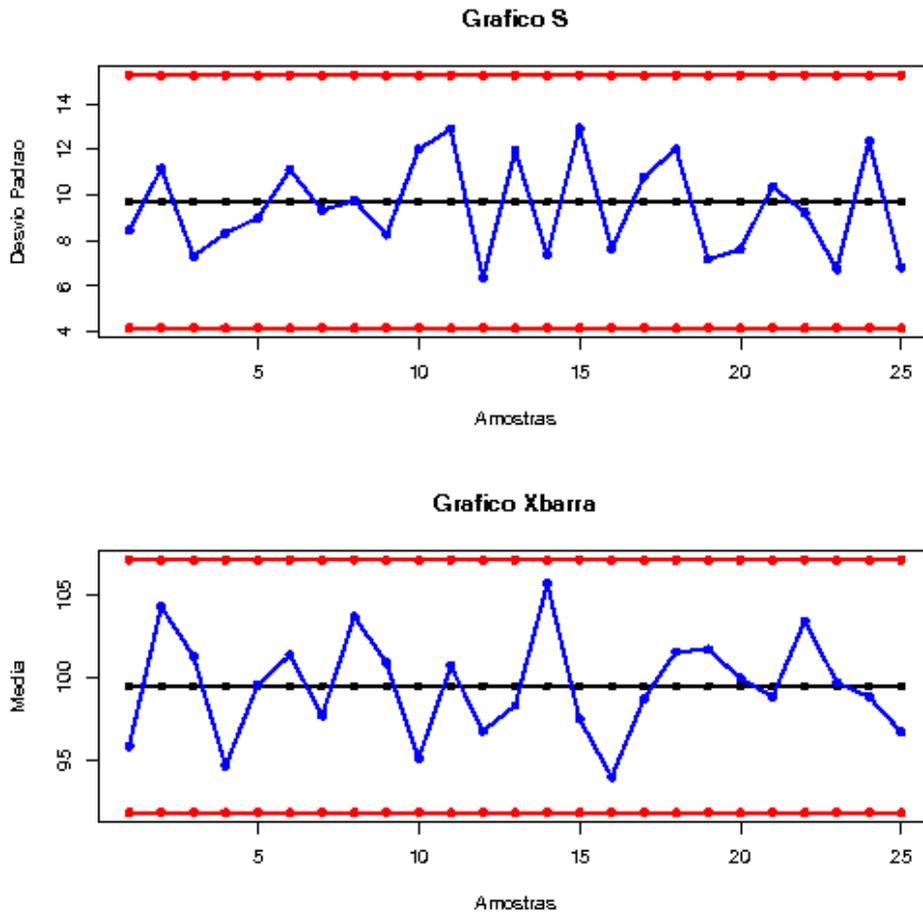


Figura 5.3.d: Relatório instantâneo gerado para os Gráficos S e \bar{X} .

Se o processo encontra-se sob controle, defina os limites de especificação para a avaliação de sua capacidade e o nível de confiança dos intervalos para cada índice.

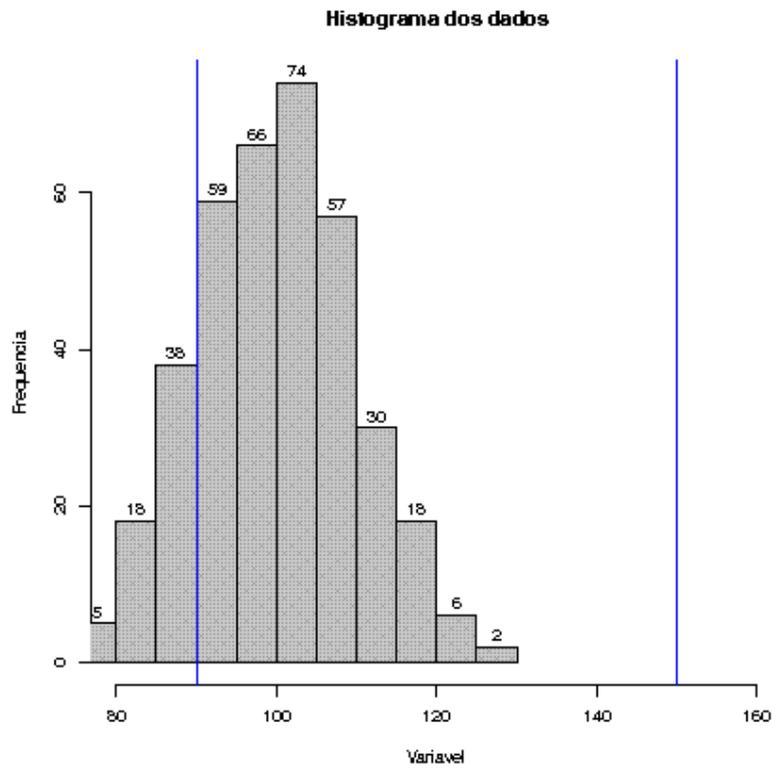
LIE:

LSE:

Alpha*:

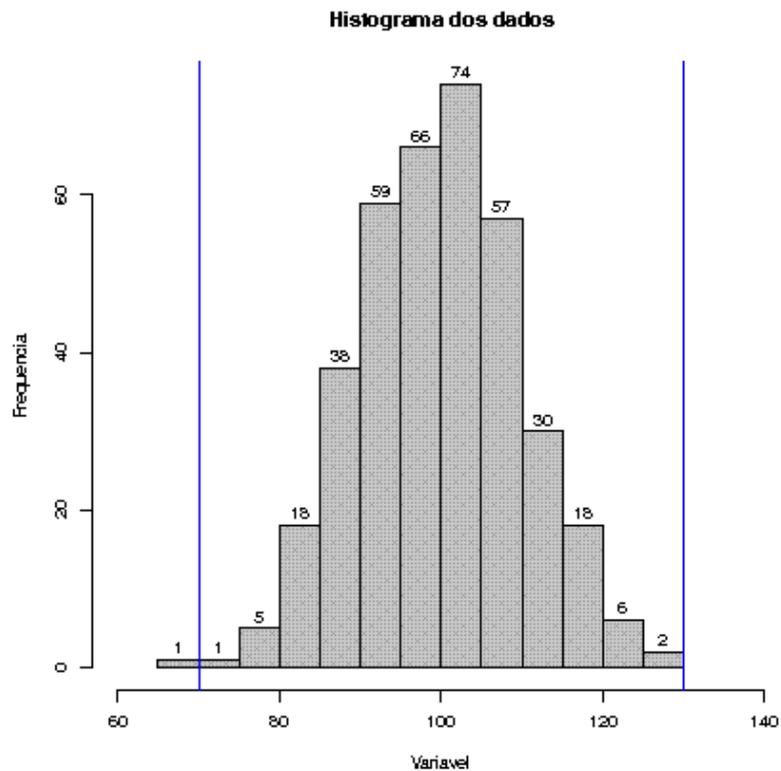
* Colocar o valor de Alpha em número decimal e usar '.' para separação das casas.

Figura 5.3.e: Definindo as condições para o estudo da capacidade do processo.



	Limite Inferior	Estimativa Pontual	Limite Superior
Cp	0.94	1.02	1.09
P (%)	91.69	98.26	105.84
Cpk	0.29	0.34	0.38
Cpm	-	0.45	-
Desvio*	-	22.37	-

Figura 5.3.f: Relatório instantâneo para a capacidade ($LIE=90$ e $LSE=150$).



	Limite Inferior	Estimativa Pontual	Limite Superior
Cp	0.94	1.02	1.09
P (%)	91.69	98.26	105.84
Cpk	0.93	1.01	1.09
Cpm	-	1.02	-
Desvio*	-	9.83	-

Figura 5.3.g: Relatório instantâneo para a capacidade ($LIE=70$ e $LSE=130$).

5.4 Considerações finais

Os procedimentos até aqui apresentados, embora sejam bastante relevantes e possam fornecer informações valiosas a respeito de um processo, foram todos aplicados à situação em que a qualidade do processo é refletida por uma única variável.

Em geral, contudo, diversas variáveis quantitativas são mensuradas em um processo e há interesse em monitorar os parâmetros de suas respectivas distribuições. A

mera aplicação dos procedimentos univariados apresentados pode conduzir a erros significativos, especialmente na medida em que o número de variáveis observadas é muito grande.

O capítulo 6 apresenta alguns dos principais procedimentos de CEP multivariado, cuja aplicação é mais adequada na grande maioria dos processos industriais monitorados.

Capítulo 6

Controle estatístico de múltiplas variáveis

É natural que a qualidade da maioria dos processos existentes seja refletida por mais do que uma variável. Em geral, o interesse está em visualizar se o comportamento de uma série de medidas em um produto apresenta um padrão homogêneo ou se por determinados períodos de tempo essa qualidade pode ficar comprometida em função de variações nas condições em que o processo ocorre.

Um exemplo que ilustra a necessidade do monitoramento simultâneo de diversas variáveis resultantes de um processo, ocorre quando se deseja monitorar as medidas de uma determinada peça de certa máquina. Em geral, devido à riqueza de detalhes associada ao projeto de um produto como este, é necessário verificar se as diversas medidas (comprimento, raio, área e volume, por exemplo) se encontram dentro de limites de variação aceitáveis, de tal forma que tal peça possa executar adequadamente suas funções. Equivalentemente, quando se deseja monitorar o padrão de qualidade de substâncias químicas ou de determinados alimentos, pode haver interesse em se monitorar as quantidades de cada componente, ingrediente ou substância presente em cada porção de itens produzidos.

Embora exista um grande potencial de aplicação para metodologias multivariadas de CEP, as metodologias univariadas são as mais conhecidas e, em diversas situações, são aplicadas de forma a substituir as metodologias multivariadas. O que ocorre é que, muitas vezes, são realizadas diversas análises univariadas do processo através de procedimentos como os descritos nos capítulo 2. De fato, os procedimentos univariados são mais fáceis de interpretar, além de possuírem a grande vantagem de já estarem disponíveis na grande maioria dos *softwares* de CEP.

Contudo, a utilização de procedimentos univariados, em detrimento dos multivariados, não se mostra uma solução adequada, uma vez que ignora a possível correlação existente entre as várias variáveis em estudo. Além disso, é possível que determinadas amostras apresentem estatísticas dentro dos limites de controle nos gráficos univariados e fora dos limites em gráficos multivariados, conduzindo assim à conclusão de que o processo se encontra sob controle, embora na verdade se encontre fora de controle. Esse tipo de conclusão incorreta é conhecido como erro tipo II. As limitações no uso de técnicas multivariadas também diminuíram sensivelmente nas

últimas décadas, visto que atualmente os computadores dispõem de elevada capacidade de processamento e tais técnicas tiveram amplo desenvolvimento e divulgação, estando acessíveis à grande maioria dos profissionais responsáveis pelo controle de qualidade dos processos.

6.1 Controle da variabilidade do processo

Em geral, o interesse inicial ao se empregar procedimentos de controle da qualidade focados no monitoramento do processo, é o de verificar se a variabilidade de um conjunto de medidas associadas a esse processo está sob controle, isto é, se a variabilidade destas medidas é constante ao longo de todo o período de tempo observado. Apenas quando se confirma que a variabilidade está sob controle é que faz sentido verificar se a média do processo também está sob controle. Sendo assim, as primeiras metodologias de controle multivariado da qualidade (CEP multivariado), apresentadas neste capítulo, são focadas no controle da variabilidade do processo (para amostras com mais de uma observação).

O primeiro procedimento discutido por Montgomery (2004, P.338 [12]) consiste em uma extensão direta do gráfico de controle univariado S^2 (detalhes a respeito da construção do Gráfico S^2 podem ser encontrados em Montgomery (2004, P.154 [12])). A estatística a ser apresentada no gráfico, para a i -ésima amostra, é dada por

$$W_i = -pn + pn \ln(n) - n \ln\left(\frac{|A_i|}{|\Sigma|}\right) + tr(\Sigma^{-1} A_i),$$

em que n é o tamanho das amostras coletadas, p é o número de variáveis observadas, tr representa o operador traço de uma matriz, Σ é a matriz de covariâncias conhecida, S_i é a matriz de covariâncias amostral da i -ésima amostra e $A_i = (n-1)S_i$. O processo é considerado fora de controle se W_i estiver acima do limite superior de controle, dado por $\chi_{p(p+1)/2;\alpha}$, que corresponde ao α -ésimo quantil de uma distribuição Qui-quadrado com $\frac{p(p+1)}{2}$ graus de liberdade.

É possível que a matriz de covariâncias, Σ , seja conhecida, com as verdadeiras variância e covariância de cada variável observada e de cada par de variáveis, respectivamente. Pode acontecer ainda que exista uma especificação que determine quais são os valores ideais para tais variâncias e covariâncias. Contudo, caso isso não

ocorra, tais valores devem ser estimados. Desta forma, a matriz de covariâncias amostral, S , pode ser utilizada como estimador de Σ , sendo tal matriz calculada a partir de amostras coletadas em análises anteriores ou mesmo através das amostras disponíveis para a análise corrente.

Um segundo procedimento de controle da variabilidade consiste em se construir o gráfico da série temporal correspondente à variância amostral generalizada de cada amostra. Dessa forma, a estatística a ser apresentada no gráfico corresponde ao determinante da matriz de covariâncias da i -ésima amostra, ou seja, $|S_i|$. Neste caso, os limites de controle são dados por:

$$|\Sigma|(b_1 \pm 3b_2^{0.5}), \quad (6.a)$$

com $b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p (n-i) \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right]$ e $b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i)$. Isso

se deve ao fato que, conforme Montgomery (2004, P.338 [12]), são válidas as relações dadas por $E(|S|) = b_1 |\Sigma|$ e $V(|S|) = b_2 |\Sigma|^2$.

A linha central do gráfico é dada por $b_1 |\Sigma|$ e o limite inferior de controle deve ser substituído por zero caso a equação deste limite, apresentada acima na equação (6.a), resulte em um número negativo. No caso em que a matriz Σ é estimada a partir da matriz de covariâncias amostral S , deve-se substituir $|\Sigma|$ por $|S|/b_1$ na equação (6.a) e no valor de referência do gráfico para se obter um estimador não-viesado do parâmetro $|\Sigma|$, isto é, um estimador que em média seja igual ao seu parâmetro. É importante ressaltar também que a literatura sugere que se utilize, em complemento a este segundo método, os gráficos individuais de controle da variabilidade (Gráficos R e S , discutidos no capítulo 2), para obter uma melhor visualização das características do processo. Isso se deve ao fato que, embora $|S|$ seja uma medida de dispersão multivariada amplamente utilizada, ainda se trata de uma representação escalar bastante simplista de um problema multivariado complexo. Pois pode acontecer, como mostra Montgomery (2004, P.339 [12]) por meio de um exemplo, que diferentes matrizes de covariâncias amostrais apresentem valores iguais para a variância amostral generalizada, embora transmitam informação consideravelmente diferente sobre a variabilidade do processo e sobre a correlação (ou covariância) entre as variáveis.

Seja qual for o método empregado, contudo, há uma distinção entre os gráficos obtidos mediante o estabelecimento de uma especificação para o processo e aqueles provenientes da estimação dos parâmetros. Quando uma especificação é dada e o processo se encontra fora de controle, conclui-se que o processo não possui as características desejadas ou que se supõe que possua. Pode ser que se imaginava ou se desejava que uma variável tivesse uma variância menor do que a que realmente tem. É possível, ou não, que exista uma mudança ao longo do tempo nos parâmetros da distribuição das variáveis observadas. Por outro lado, só se pode confirmar se tal mudança ocorreu, de fato, através dos gráficos de controle obtidos por meio da estimação de tais parâmetros (pois esta é a atribuição de tais gráficos).

Um algoritmo para construção e análise de gráficos de controle da variabilidade para processos multivariados, é apresentado no Quadro 6.1.a do Apêndice C.

6.2 Gráfico T^2 de Hotelling

O Gráfico T^2 de Hotelling é a extensão multivariada do Gráfico \bar{X} , discutido em detalhes no capítulo 2, e consiste no procedimento mais utilizado para monitorar o vetor de médias do processo. É um gráfico útil especialmente quando há necessidade de se identificar grandes mudanças na média do processo. Sua grande vantagem é que permite que toda a informação presente nas amostras disponíveis seja resumida em uma única estatística que segue uma distribuição conhecida. Isso se mostra útil devido ao número elevado de variáveis observadas.

Para demonstrar a utilidade desse gráfico, pode-se introduzir uma situação simples, em que são observadas apenas três variáveis quantitativas referentes a certo processo. Suponha, por exemplo, que são registradas três medidas da qualidade em um determinado tipo de suco: a quantidade de açúcar, a acidez e a quantidade de água presentes em uma porção de 100 ml. Dessa forma, a representação das médias da i -ésima amostra observada exige que se utilizem três dimensões. Contudo, para que seja preservada a informação referente à ordem de coleta dos dados, faz-se necessário o uso de uma quarta dimensão, o que já torna demasiadamente complexa a tarefa de representar graficamente os resultados observados nas amostras. A dificuldade aumenta ainda mais à medida em que aumenta o número de variáveis observadas. Contudo,

conforme é discutido nas seções 6.2.1 e 6.2.2, a utilização do Gráfico T^2 permite a representação, em apenas duas dimensões, dos resultados observados nas amostras, sem grandes perdas de informação.

6.2.1 Análise com dados agrupados

Assim como discutido no capítulo 2, em diversas situações há interesse em verificar se o processo é capaz de atender às especificações estabelecidas *a priori* para a média e para a variabilidade de um conjunto de variáveis. Tais especificações são estabelecidas de acordo com as necessidades e desejos do consumidor, as exigências de órgãos de fiscalização do governo ou de acordo com critérios referentes ao projeto do produto que garantam o bom funcionamento do item produzido.

Para se introduzir o Gráfico T^2 para dados agrupados, inicialmente define-se $\mathbf{x}'_{ij} = (x_{1ij}, x_{2ij}, \dots, x_{pij})$ como sendo o vetor de variáveis de interesse referente à j -ésima observação da i -ésima amostra, para $i=1,2,\dots,m$ e $j=1,2,\dots,n$. Para se construir esse gráfico, supõe-se que as p variáveis observadas, isto é, $\mathbf{x}' = (x_1, x_2, \dots, x_p)$, sigam, aproximadamente, uma distribuição Normal p -variada com vetor de médias $\boldsymbol{\mu}' = (\mu_1, \mu_2, \dots, \mu_p)$ e matriz de covariâncias Σ , cuja função de densidade de probabilidade multivariada é dada por:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \quad (6.2.1.a)$$

em que $-\infty < x_k < +\infty$, $k=1,2,\dots,p$. Outra definição importante é dada por:

$$\bar{\mathbf{x}}_i = \begin{pmatrix} \bar{x}_{1i} \\ \bar{x}_{2i} \\ \vdots \\ \bar{x}_{pi} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} \sum_{j=1}^n x_{1ij} \\ \frac{1}{n} \sum_{j=1}^n x_{2ij} \\ \vdots \\ \frac{1}{n} \sum_{j=1}^n x_{pij} \end{pmatrix}, \quad (6.2.1.b)$$

que corresponde ao vetor de médias da i -ésima amostra. O Gráfico T^2 de Hotelling é a representação da série temporal da estatística T^2 , definida para a i -ésima amostra, como:

$$T_i^2 = n(\bar{\mathbf{x}}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}}_i - \boldsymbol{\mu}), \quad (6.2.1.c)$$

em que $\boldsymbol{\mu}$ é o vetor de médias estabelecido como ideal para o processo e $\boldsymbol{\Sigma}$ é a matriz de covariâncias das variáveis observadas. O limite superior de controle desse gráfico, em função de que os dados seguem uma distribuição Normal Multivariada, é dado pelo α -ésimo quantil de uma distribuição Qui-quadrado com p graus de liberdade, $\chi_{p;\alpha}^2$. O Gráfico T^2 não apresenta valor de referência ou linha central e o seu limite inferior de controle é sempre igual a zero.

Contudo, é comum também que não seja dada qualquer especificação para o processo. Neste caso, os parâmetros da distribuição multivariada das variáveis observadas devem ser estimados a partir de todas as amostras disponíveis. Tais estimativas, no caso do vetor de médias, podem ser dadas por:

$$\bar{\bar{\mathbf{x}}} = \begin{pmatrix} \bar{\bar{x}}_1 \\ \bar{\bar{x}}_2 \\ \vdots \\ \bar{\bar{x}}_p \end{pmatrix} = \begin{pmatrix} \frac{1}{m} \sum_{i=1}^m \bar{x}_{1i} \\ \frac{1}{m} \sum_{i=1}^m \bar{x}_{2i} \\ \vdots \\ \frac{1}{m} \sum_{i=1}^m \bar{x}_{pi} \end{pmatrix}. \quad (6.2.1.d)$$

Observe que $\bar{\bar{\mathbf{x}}}$ é um estimador não-viesado de $\boldsymbol{\mu}$ quando o processo se encontra sob

controle, pois $E(\bar{\bar{x}}_k) = E\left(\frac{1}{m} \sum_{i=1}^m \bar{x}_{ki}\right) = E\left(\frac{1}{n} \sum_{j=1}^n x_{kij}\right) = E(x_{kij}) = \mu_k$, para $k=1,2,\dots,p$.

Caso o processo se encontre sob controle, um estimador não-viesado da matriz de covariâncias $\boldsymbol{\Sigma}$ é dado pela média S de dimensão $p \times p$ das matrizes de covariâncias amostrais, escrita como:

$$S = \begin{bmatrix} \bar{S}_1^2 & \bar{S}_{12} & \dots & \bar{S}_{1p} \\ \bar{S}_{12} & \bar{S}_2^2 & \dots & \bar{S}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{S}_{1p} & \bar{S}_{2p} & \dots & \bar{S}_p^2 \end{bmatrix}. \quad (6.2.1.e)$$

Os elementos de S são dados por:

$$\left\{ \begin{array}{l} \bar{S}_k^2 = \frac{1}{m} \sum_{i=1}^m S_{ki}^2 = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{(n-1)} \sum_{j=1}^n (x_{kij} - \bar{x}_{ki})^2 \right) = \frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (x_{kij} - \bar{x}_{ki})^2 \\ \bar{S}_{kg} = \frac{1}{m} \sum_{i=1}^m S_{kgi} = \frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (x_{kij} - \bar{x}_{ki})(x_{gij} - \bar{x}_{gi}) \end{array} \right. , \quad (6.2.1.f)$$

para $k, g=1,2,\dots,p$ e $k \neq g$.

Assim, dispondo dessas estimativas para μ e Σ , a estatística a ser apresentada no Gráfico T^2 quando não são estabelecidas especificações para o processo, é dada por:

$$T_i^2 = n(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' S^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}). \quad (6.2.1.g)$$

Alt (1985 [1]) destaca que, ao realizar o cálculo do limite superior de controle a partir das amostras disponíveis atualmente, para o monitoramento futuro do processo por meio de novas amostras, existem duas diferentes etapas as quais exigem diferentes limites de controle. Inicialmente, deve-se determinar quais são as amostras que se encontram sob controle, a fim de selecionar uma amostra que represente adequadamente os parâmetros da distribuição das p variáveis observadas. Após a seleção de tais amostras, é possível obter estimativas mais confiáveis para μ e Σ , conduzindo, com isso, a certa alteração no valor do limite superior de controle que deve guiar o monitoramento futuro do processo.

Na etapa inicial da análise (ou etapa 1), o limite superior de controle é dado por:

$$LSC = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{p, mn-m-p+1; \alpha}, \quad (6.2.1.h)$$

sendo que $F_{p, mn-m-p+1; \alpha}$ é o α -ésimo quantil de uma distribuição F com p graus de liberdade no numerador e $mn-m-p+1$ graus de liberdade no denominador. Na segunda etapa da análise (ou etapa 2), este limite é dado por:

$$LSC = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{p, mn-m-p+1; \alpha}, \quad (6.2.1.i)$$

O Quadro 6.2.a do Apêndice C apresenta o algoritmo para construção e análise do Gráfico T^2 , no caso em que são coletadas amostras com mais de uma observação cada (dados agrupados).

6.2.2 Análise com dados individuais

Existem situações em que se dispõe de amostras de tamanho um, ao mesmo tempo em que diversas características devem ser monitoradas no controle de qualidade do processo. O Gráfico T^2 referente a esta situação consiste de um caso particular daquela descrita na seção 6.2.1. Dessa forma, quando são estabelecidas especificações para o processo, a estatística apresentada no Gráfico T^2 continua sendo a apresentada na equação (6.2.1.c) e o limite superior de controle ainda é o α -ésimo quantil de uma distribuição Qui-quadrado com p graus de liberdade.

Contudo, quando não são estabelecidas especificações, a simplificação da notação mostra-se conveniente. Portanto, a variável observada passa a ser definida como x_{ki} , correspondente à i -ésima observação (ou amostra) da k -ésima variável. O vetor de médias do processo passa então a ser estimado através de:

$$\bar{\mathbf{x}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix} = \begin{pmatrix} \frac{1}{m} \sum_{i=1}^m x_{1i} \\ \frac{1}{m} \sum_{i=1}^m x_{2i} \\ \vdots \\ \frac{1}{m} \sum_{i=1}^m x_{pi} \end{pmatrix}. \quad (6.2.2.a)$$

A estatística T^2 de Hotelling a ser apresentada no gráfico, para o caso de observações individuais, torna-se então igual a:

$$T_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})' S^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}). \quad (6.2.2.b)$$

Na etapa inicial (etapa 1) de construção do limite superior de controle, em que há interesse apenas em identificar quais são as amostras que se encontram sob controle no conjunto de dados disponível, o limite superior de controle é dado por:

$$LSC = \frac{(m-1)^2}{m} \beta_{p/2, (m-p-1)/2; \alpha} \quad (6.2.2.c)$$

(Tracy, Young e Mason, 1992 [21]), em que $\beta_{p/2, (m-p-1)/2; \alpha}$ é o α -ésimo quantil de uma distribuição Beta com parâmetros $p/2$ e $(m-p-1)/2$. O limite superior de controle da segunda etapa da análise (etapa 2), em que são estabelecidos os limites de controle a serem utilizados em análises futuras, é dado por:

$$LSC = \frac{p(m+1)(m-1)}{m^2 - mp} F_{p, m-p; \alpha}. \quad (6.2.2.d)$$

Contudo, no caso em que se coletam observações individuais do vetor de variáveis de interesse, ou seja, amostras de tamanho 1, faz-se necessário o uso de estimadores mais sofisticados da matriz de covariâncias Σ . Sullivan e Woodall (1995 [17]) discutem esse problema e comparam vários estimadores que poderiam ser empregados neste caso. Um desses estimadores é o “usual” obtido pela simples combinação das m observações disponíveis e é dado por:

$$S_1 = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'. \quad (6.2.2.e)$$

Um outro estimador, sugerido originalmente por Holmes e Mergen (1993 [8]), considera a diferença entre pares consecutivos de observações, ou seja, $\mathbf{v}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$, para $i=1,2,\dots,m-1$. Tais elementos são arranjados em uma matriz, de forma que é constituída a matriz dada por:

$$V = \begin{bmatrix} \mathbf{v}'_1 \\ \mathbf{v}'_2 \\ \dots \\ \mathbf{v}'_{m-1} \end{bmatrix}. \quad (6.2.2.f)$$

Finalmente, estima-se Σ através de:

$$S_2 = \frac{1}{2} \frac{V'V}{(m-1)}, \quad (6.2.2.g)$$

que é a metade da matriz de covariâncias amostral dessas diferenças.

Um algoritmo de construção e análise do Gráfico T^2 para amostras com apenas uma observação (dados individuais), é apresentado no Quadro 6.2.b do Apêndice C.

6.3 Exemplo 12

O primeiro conjunto de dados analisado neste capítulo está exposto na Tabela 6.3 do Apêndice B e consiste de 300 observações divididas em 25 amostras de mesmo tamanho. São observadas 3 variáveis em cada unidade amostral, cada uma corresponde a uma medida quantitativa referente aos aparelhos celulares produzidos pela empresa fictícia responsável pelos exemplos dos capítulos anteriores. Tais medidas poderiam ser,

por exemplo, o comprimento, a largura e a altura, todas em milímetros (mm), de determinada peça interna desses aparelhos. Suponha que o interesse da empresa é manter o processo sob controle, mas esta não tem especificações em relação aos parâmetros (vetor de médias e matriz de covariâncias) da distribuição multivariada das variáveis observadas.

A seguir, nas Figuras 6.3.a, 6.3.b, 6.3.c e 6.3.d é apresentada uma sequência de passos para a análise do conjunto de dados em questão no sistema de CEP *on-line* proposto, desde a inserção dos dados (Figura 6.3.a) até a visualização do relatório instantâneo para os gráficos de controle da variabilidade e Gráfico T^2 (Figura 6.3.d).

Observa-se na Figura 6.3.c, que foi selecionada a opção “Etapa 1”, já que o interesse inicial da análise reside em identificar as amostras que estão fora de controle.

Na Figura 6.3.d são apresentados, primeiramente, os gráficos de controle da variabilidade e, em seguida, o Gráfico T^2 . Os gráficos de controle da variabilidade foram construídos segundo os dois procedimentos discutidos na seção 6.1. O gráfico à esquerda corresponde ao primeiro procedimento de controle da variabilidade discutido, e o da direita, ao segundo. Ambos não detectaram sinal de falta de controle do processo com respeito à variabilidade, uma vez que não foram observados pontos que excedam o limite superior de controle. No entanto, o Gráfico T^2 identificou as 5 últimas amostras como fora de controle com relação à média, visto que tais amostras estão todas localizadas acima do limite superior de controle. Portanto, a análise deve prosseguir, na tentativa de identificar as causas responsáveis pela falta de controle das 5 últimas amostras observadas. Se tais causas forem identificadas, deve-se ainda retirar essas amostras fora de controle do banco de dados e reconstruir os gráficos (tais gráficos não são exibidos aqui) a partir das 20 amostras restantes. Caso não seja identificada mais nenhuma amostra fora de controle com relação à variabilidade ou à média, pode-se ainda estimar o vetor de médias, a matriz de covariâncias e o limite superior de controle (também não são exibidos aqui), a serem utilizados no monitoramento posterior do processo.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (n é constante):

Defina o nome da sua tabela de dados:

Defina o número de variáveis do processo:

Cole aqui os dados referentes a cada variável, sendo a primeira linha o nome da variável:

v1	v2	v3
V1	V2	V3
100.84	99.29	99.19
99.15	98.96	100.35
98.25	99.63	100.17
98.31	100.83	97.84
99.99	100.50	101.34
100.04	99.77	101.77
97.94	99.50	98.64
101.96	100.41	99.60
99.85	99.95	99.77
99.89	100.43	100.36
101.31	101.11	102.08
100.23	99.07	101.93

Avançar

Figura 6.3.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo12
Número de variáveis:	3
Número de observações:	300
Número de amostras:	25
Tamanho das amostras:	12

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 6.3.b: Confirmando os dados inseridos.

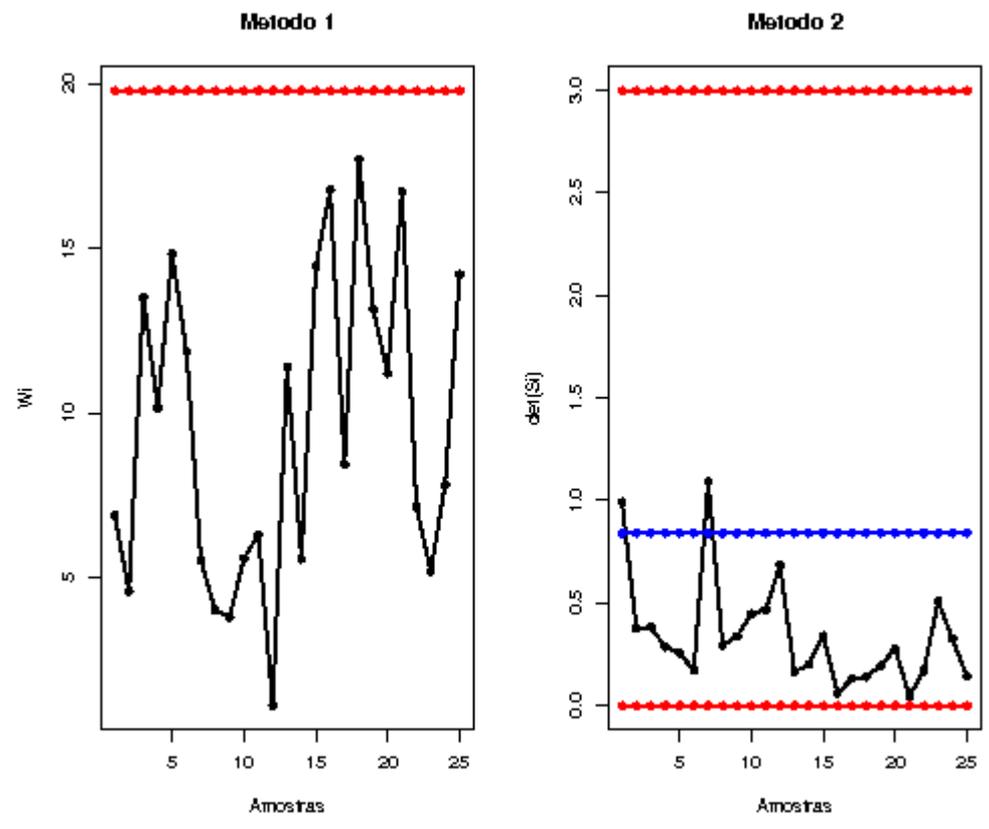
Defina as condições para os gráficos:

Defina a etapa atual da análise*:

* A Primeira Etapa (ou Etapa 1) da análise visa a identificação das amostras fora de controle pelo gráfico T2, enquanto a Segunda Etapa (ou Etapa 2) visa o monitoramento futuro do processo, quando retiradas todas as amostras fora de controle, com a obtenção de estimativas confiáveis do vetor de médias, da matriz de covariâncias e do limite superior de controle.

Avançar

Figura 6.3.c: Definindo as características e condições do processo.



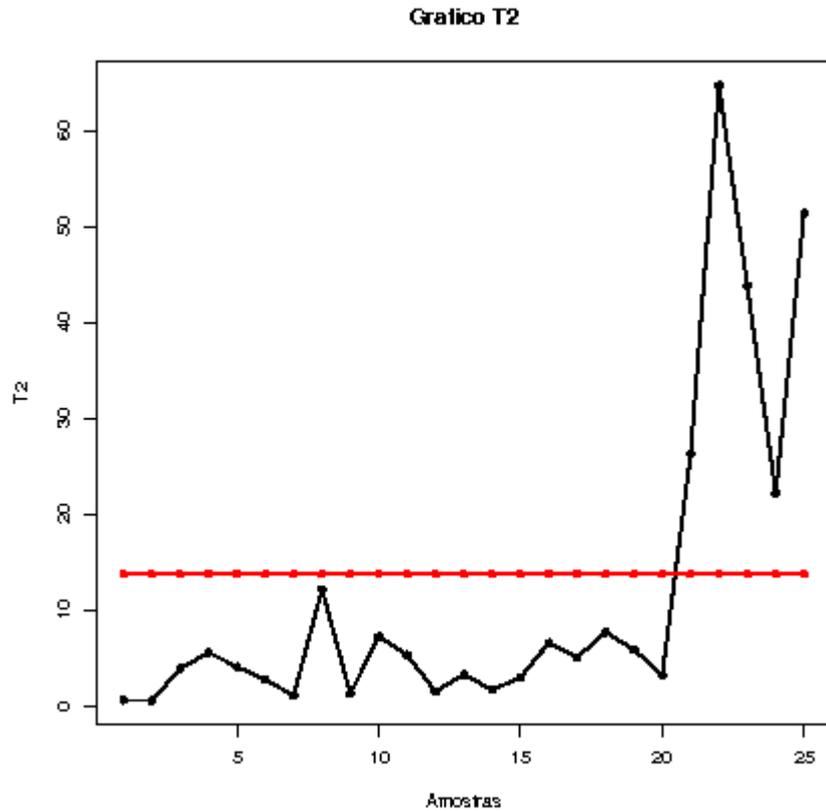


Figura 6.3.d: Relatório instantâneo gerado para os gráficos de controle da variabilidade e Gráfico T^2 (dados agrupados).

6.4 Exemplo 13

O segundo conjunto de dados a ser analisado neste capítulo está exposto na Tabela 6.4 do Apêndice B e novamente refere-se a observações das 3 medidas quantitativas de determinada peça interna dos aparelhos celulares produzidos por certa empresa de telecomunicações (fictícia). Tal conjunto é constituído por 25 amostras com apenas uma observação cada. Suponha que o interesse da empresa é manter o processo sob controle, considerando as seguintes especificações para o vetor de médias e para a matriz de covariâncias:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 100 \end{pmatrix} \quad \text{e} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0,5 & 0,5 \\ 0,5 & 1 & 0,5 \\ 0,5 & 0,5 & 1 \end{bmatrix}.$$

A seguir, nas Figuras 6.4.a, 6.4.b, 6.4.c e 6.4.d são apresentados os passos necessários para a análise do conjunto de dados em questão no sistema de CEP *on-line*

desenvolvido, desde a inserção dos dados (Figura 6.4.a) até a visualização do relatório instantâneo para o Gráfico T^2 (Figura 6.4.d).

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (n é constante):

Defina o nome da sua tabela de dados:

Defina o número de variáveis do processo:

Cole aqui os dados referentes a cada variável, sendo a primeira linha o nome da variável:

v1	v2	v3
v1	v2	v3
98.37	99.56	97.46
100.56	100.07	100.53
101.24	99.76	100.73
99.21	100.95	100.62
99.05	99.84	99.42
99.78	99.52	99.37
98.96	99.64	100.00
99.65	99.90	99.01
98.05	99.89	99.97
98.68	101.08	98.70
100.53	100.33	100.64
99.08	99.42	98.30

Figura 6.4.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo13
Número de variáveis:	3
Número de observações:	25
Número de amostras:	25
Tamanho das amostras:	1

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Figura 6.4.b: Confirmando os dados inseridos.

Defina as condições para os gráficos:

Médias das variáveis:

V1	<input type="text" value="100"/>	V2	<input type="text" value="100"/>	V3	<input type="text" value="100"/>
----	----------------------------------	----	----------------------------------	----	----------------------------------

Variâncias e Covariâncias entre as variáveis:

	V1	V2	V3
V1	<input type="text" value="1"/>	<input type="text" value="0.5"/>	<input type="text" value="0.5"/>
V2	<input type="text" value="0.5"/>	<input type="text" value="1"/>	<input type="text" value="0.5"/>
V3	<input type="text" value="0.5"/>	<input type="text" value="0.5"/>	<input type="text" value="1"/>

Figura 6.4.c: Definindo as condições do processo.

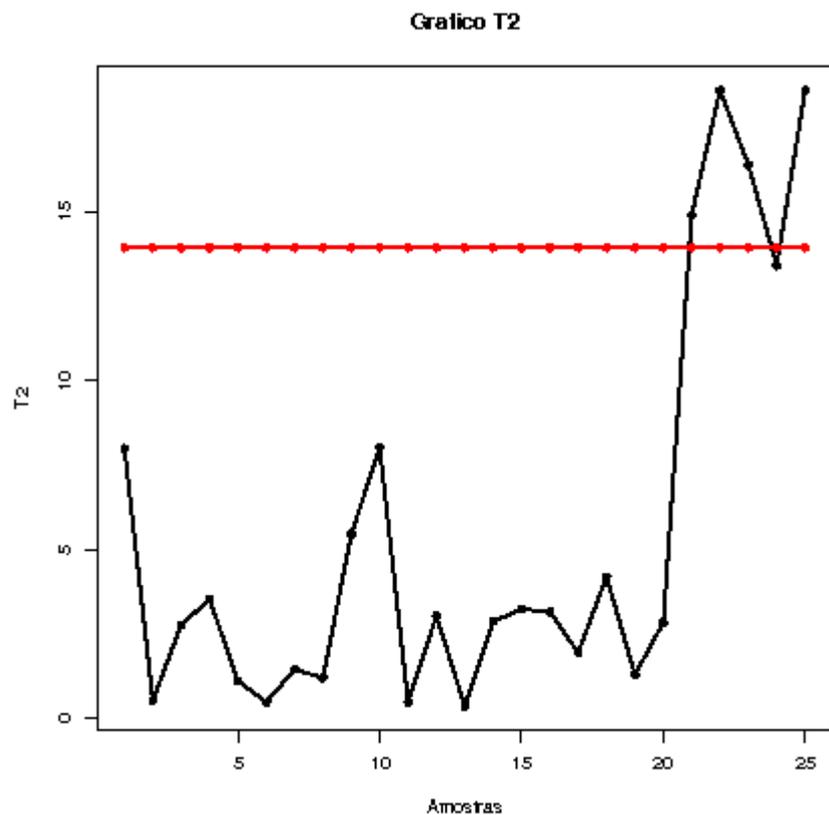


Figura 6.4.d: Relatório instantâneo gerado para o Gráfico T^2 (dados individuais).

O Gráfico T^2 apresentado na Figura 6.4.d detectou a falta de controle das 5 últimas amostras, pois tais amostras, com exceção da amostra de número 24, estão localizadas acima do limite superior de controle. Para o caso em que não se consideram as especificações dos parâmetros (Figura 6.4.e), o comportamento do Gráfico T^2 é equivalente ao daquele apresentado na Figura 6.4.d apenas quando se considera o segundo estimador da matriz de covariâncias discutido na seção 6.2.2 (equação (6.2.2.g)). Em contrapartida, o Gráfico T^2 proveniente do primeiro estimador da matriz de covariâncias discutido (equação (6.2.2.e)), apresentado na Figura 6.4.e (gráfico à esquerda), não foi capaz de identificar a falta de controle no processo para este conjunto de dados especificamente.

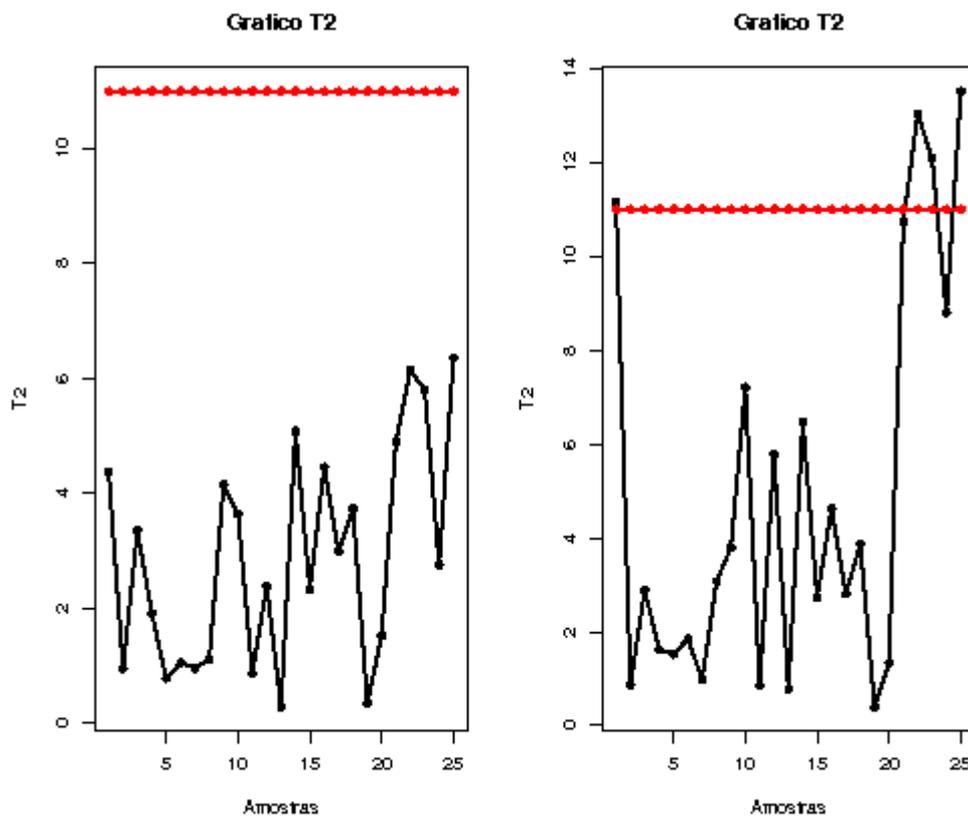


Figura 6.4.e: Relatório instantâneo gerado para o Gráfico T^2 sem especificação (dados individuais).

Como a falta de controle do processo é diagnosticada através de dois dos três gráficos construídos, é importante investigar se houve alguma alteração nas condições de ocorrência do processo durante o período em que as 5 últimas amostras foram produzidas, visando determinar se alguma dessas alterações pode ser considerada como

responsável pelo comportamento atípico em pelo menos uma das 3 medidas das peças internas dos aparelhos produzidos, o que caracterizaria a condição do processo como fora de controle.

6.5 Considerações finais

As metodologias abordadas neste capítulo permitem o monitoramento das médias e da variabilidade de um processo em que se observam várias variáveis quantitativas capazes de refletir o nível de qualidade do processo. Tais metodologias permitem que esse monitoramento seja realizado de uma forma na qual a estrutura de associação entre as variáveis não é ignorada, o que por sua vez conduz a resultados mais precisos e coerentes com a realidade do fenômeno em estudo.

Diversas outras metodologias têm sido desenvolvidas e incorporadas ao conjunto de ferramentas disponíveis nos principais *softwares* de CEP. Entre tais, encontram-se extensões multivariadas do Gráfico de Somas Acumuladas (ou Cusum) e do Gráfico de Médias Móveis Exponencialmente Ponderadas (ou MMEP), que podem ser utilizados para identificar mudanças mais sutis nos parâmetros do processo. Os métodos baseados em regressão podem ser utilizados para monitorar um processo no qual existem variáveis referentes às condições do processo e dos insumos utilizados (independentes ou explicativas) e variáveis referentes à qualidade do item produzido (respostas). Neste caso, as variáveis independentes são utilizadas (controladas) de forma a diminuir a variabilidade no resultado das respostas, promovendo assim resultados mais precisos. Além disso, os métodos baseados em análise discriminante consistem em outra possibilidade para o monitoramento do processo.

Em relação à avaliação da capacidade de processos multivariados, encontram-se na literatura propostas que vão desde índices calculados separadamente para cada característica de qualidade, até índices mais elaborados que tentam, de alguma forma, compôr a informação geral de todas as variáveis (Taan, Subbaiah et. al, 1993 [18]; Zhang, 1998 [28]). No próximo capítulo são discutidos alguns índices de capacidade para dados de processos normais multivariados, que podem ser obtidos usando a técnica de Análise de Componentes Principais.

Capítulo 7

Capacidade do processo multivariado

Ultimamente tem-se discutido muito sobre como deveria ser realizado o estudo da capacidade de um processo com múltiplas características da qualidade. Uma estratégia, a princípio, seria analisar a capacidade do processo em cada variável em questão, por meio dos índices descritos no capítulo 5. Contudo, este procedimento não é o mais recomendável, visto que não incorpora a possível correlação existente entre as várias características de qualidade.

Vários autores têm proposto definições alternativas de Índices de Capacidade de Processos Multivariados (ICPMs), baseando-se em diferentes aproximações. Em geral, ICPMs podem ser obtidos (construídos) usando:

1. A razão entre uma região de especificação (ou de tolerância) e uma região de processo. Esta aproximação é semelhante à dos índices de capacidade univariados convencionais, tal como o $C_p = (LSE - LIE) / 6\sigma$. Uma discussão detalhada a respeito dos índices obtidos a partir deste tipo de aproximação pode ser encontrada em Pearn, Kotz e Johnson (1992 [13]), Kotz e Johnson (1993 [10]), Chan, Cheng e Spiring (1991 [5]).
2. A probabilidade do item defeituoso. Neste caso, a função de distribuição de probabilidade multivariada é usada para calcular a probabilidade de um item qualquer produzido ser defeituoso (Wierda, 1993 [27]; Chen, 1994 [6]).
3. A Análise de Componentes Principais (Wang e Chen, 1998 [24]; Wang e Du, 2000 [25]; Veevers, 1999 [22]; Shinde e Khadse, 2008 [15]).
4. Outras aproximações baseadas em funções de perda (Pearn, Kotz e Johnson, 1992 [13]) ou representação de vetor (Shahriari, Hubele e Lawrence, 1995 [14]).

Neste capítulo são abordados somente os ICPMs que podem ser obtidos usando a técnica estatística multivariada denominada Análise de Componentes Principais (descrita em detalhes na seção 7.1), quando é plausível supor que os dados coletados sejam oriundos de processos normais multivariados. Esta condição fundamental para o uso dos ICPMs discutidos pode ser verificada usando uma generalização do teste de

Shapiro-Wilk para normalidade multivariada, proposta por Villasenor-Alva e Gonzalez-Estrada (2009 [23]).

7.1 Análise de Componentes Principais

Um componente principal (CP) é uma combinação linear das variáveis originalmente observadas, X_1, X_2, \dots, X_p , que explica o máximo possível da variabilidade dos dados, variabilidade esta que não tenha sido explicada pelos demais componentes. Ou seja, os diversos componentes principais são independentes entre si e cada um deles contribui de uma forma particular ao explicar as variáveis (ou dados) observadas. Em termos matemáticos, tem-se que:

$$\begin{aligned} Z_1 &= a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Z_2 &= a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Z_p &= a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p, \end{aligned} \tag{7.1.a}$$

em que Z_i é o i -ésimo componente principal, X_i é a i -ésima variável originalmente observada e a_{ij} é uma constante a ser determinada. Seja $\mathbf{X}' = (X_1, X_2, \dots, X_p)$ um vetor aleatório com matriz de covariâncias Σ , e sejam $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ os autovalores de Σ . A constante a_{ij} corresponde ao j -ésimo elemento do autovetor associado ao i -ésimo autovalor λ_i .

A interpretação de cada componente é condicionada ao valor das constantes a_{ij} , também denominadas cargas. Se um componente possui algumas dessas constantes bem próximas de zero, outras consideravelmente maiores que zero e ainda outras consideravelmente menores que zero, a interpretação desse componente é então no sentido de contraste entre um melhor desempenho nas variáveis cujas constantes são maiores que zero e um melhor desempenho naquelas que apresentam constantes menores que zero, sendo que as variáveis com coeficientes próximos de zero não são bem explicadas por esse componente. Por exemplo, ao se constatar que uma observação possui um escore negativo de alta magnitude em um determinado componente, pode-se concluir que esta observação apresenta valores relativamente altos em uma ou mais das

variáveis cujos coeficientes são consideravelmente menores do que zero neste componente.

Muitas vezes os componentes principais são calculados a partir de variáveis padronizadas, isto é, com média igual a 0 e variância igual a 1. Uma razão para esta prática comum está no fato de que, em geral, as variáveis originais do processo são expressas em grandezas diferentes, podendo, assim, apresentar resultados de magnitudes bem diferentes. Dessa forma, uma variável pode parecer contribuir muito para a variabilidade total do processo, apenas porque está em uma escala de medida de maior magnitude que a das demais variáveis. A padronização das variáveis resolve este tipo de problema. E nestes casos, a matriz de covariâncias Σ corresponde à matriz de correlações das variáveis originalmente observadas.

A variância do i -ésimo componente principal é o i -ésimo autovalor λ_i . Como o conjunto completo dos p componentes principais apresenta a mesma informação contida nas p variáveis originalmente observadas, então a proporção da variabilidade dos dados que é bem explicada pelo i -ésimo componente, é dada por:

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, \quad (7.1.b)$$

A partir dessa proporção e do número de variáveis que são bem explicadas pelos primeiros componentes é que se determina o número ideal, v , de dimensões para a análise. Em geral, há interesse em se utilizar o menor número possível de componentes principais, que permita que um número relativamente grande das variáveis observadas e uma proporção razoável (em torno de 0,9 ou 90%) da variabilidade dos dados sejam bem explicados. Portanto, é possível reduzir a dimensionalidade do problema em estudo através do uso da técnica de Análise de Componentes Principais, pois tal técnica permite que as p variáveis originalmente observadas sejam substituídas, na análise, pelos v ($v < p$) primeiros componentes principais, sem grandes perdas de informação.

7.2 Dados normais multivariados

Se as características de qualidade observadas são provenientes de uma distribuição normal multivariada, então a Análise de Componentes Principais (ACP) pode ser aplicada para o estudo da capacidade do processo. Conseqüentemente, as novas

variáveis obtidas, isto é, os componentes principais (CPs) são mutuamente independentes e normalmente distribuídos (Tong, 1990 [20]). Nas seções seguintes são discutidos alguns dos principais índices de capacidade para processos normais multivariados (ICPMs) encontrados na literatura, dentre os quais, os ICPMs definidos por Wang e Chen (1998 [24]), que utilizam os índices univariados C_p , C_{pk} e C_{pm} calculados para os primeiros v CPs; o ICPM proposto por Veevers (1999 [22]), que se baseia apenas no primeiro CP; e por último, um método alternativo, proposto por Shinde e Khadse (2008 [15]), para estimar a capacidade de processos normais multivariados a partir da distribuição de probabilidade empírica dos CPs (que também seguem distribuição normal multivariada). Tal método foi proposto como uma alternativa aos outros índices mencionados, devido a problemas observados quando da estimação da capacidade de processos multivariados por meio destes índices, bem como problemas existentes em sua própria definição. Tais problemas são comentados nas seções que seguem.

7.2.1 ICPM-I

O índice definido por Wang e Chen (1998 [24]) para estimar a capacidade de processos normais multivariados, é dado por:

$$MC_p = \left(\prod_{i=1}^v C_{p;CP_i} \right)^{1/v}, \quad (7.2.1.a)$$

em que $C_{p;CP_i} = (LSE_{CP_i} - LIE_{CP_i}) / 6\sigma_{CP_i}$ representa a medida univariada da capacidade do processo para o i -ésimo CP e v denota o número de CPs que, juntos, explicam aproximadamente 90% da variabilidade do processo. Os limites de especificação dos CPs, bem como seus valores alvo, são dados por:

$$LIE_{CP_i} = \mathbf{u}'_i \mathbf{LIE}, \quad LSE_{CP_i} = \mathbf{u}'_i \mathbf{LSE}, \quad T_{CP_i} = \mathbf{u}'_i \mathbf{T}, \quad (7.2.1.b)$$

sendo que $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ são os autovetores da matriz de covariâncias Σ , p é o número de variáveis do processo e os vetores \mathbf{LIE} , \mathbf{LSE} e \mathbf{T} representam, respectivamente, os limites inferiores de especificação, os limites superiores de especificação e os valores-alvo das variáveis originalmente observadas, $\mathbf{X}' = (X_1, X_2, \dots, X_p)$.

A interpretação do MC_p é análoga à do C_p , ou seja, valores maiores do que 1 são desejados e indicam que o processo dispõe de condições para atender às especificações estabelecidas.

Contudo, muitas vezes é necessário estimar o valor de σ_{CP_i} , obtendo-se uma estimativa de $C_{p;CP_i}$, $\hat{C}_{p;CP_i} = (LSE_{CP_i} - LIE_{CP_i}) / 6S_{CP_i}$, e conseqüentemente, uma estimativa de MC_p ,

$$M\hat{C}_p = \left(\prod_{i=1}^v \hat{C}_{p;CP_i} \right)^{1/v}. \quad (7.2.1.c)$$

Como os CPs são normalmente distribuídos, então um intervalo de confiança de nível aproximado $100(\alpha)\%$ para MC_p , é dado por:

$$\left[\left(\prod_{i=1}^v \hat{C}_{p;CP_i} \sqrt{\frac{\chi_{n-1;(1-\alpha)/2}^2}{n-1}} \right)^{1/v}; \left(\prod_{i=1}^v \hat{C}_{p;CP_i} \sqrt{\frac{\chi_{n-1;(1+\alpha)/2}^2}{n-1}} \right)^{1/v} \right] \quad (7.2.1.d)$$

(Kushler e Hurley, 1992 [11]), com $\chi_{n-1;(1-\alpha)/2}^2$ e $\chi_{n-1;(1+\alpha)/2}^2$ representando, respectivamente, o $100(1-\alpha)/2$ -ésimo e o $100(1+\alpha)/2$ -ésimo quantil de uma distribuição Qui-quadrado com $n-1$ graus de liberdade.

De maneira análoga, Wang e Chen (1998 [24]) definiram os índices MC_{pk} e MC_{pm} , ao substituir $C_{p;CP_i}$ na equação (7.2.1.a) por $C_{pk;CP_i}$ e $C_{pm;CP_i}$, respectivamente.

No caso, $\hat{C}_{pk;CP_i} = \min \left(\frac{LSE_{CP_i} - \bar{x}_{CP_i}}{3S_{CP_i}}; \frac{\bar{x}_{CP_i} - LIE_{CP_i}}{3S_{CP_i}} \right)$ e $\hat{C}_{pm;CP_i} = \frac{\hat{C}_{p;CP_i}}{\sqrt{1+V_{CP_i}^2}}$, com

$V_{CP_i}^2 = \frac{\bar{x}_{CP_i} - T_{CP_i}}{S_{CP_i}}$. Um intervalo de confiança para MC_{pk} , ao nível aproximado de

$100(\alpha)\%$, é dado por:

$$\left[\left(\prod_{i=1}^v \hat{C}_{pk;CP_i} \left[1 - Z_{(1+\alpha)/2} \sqrt{\frac{1}{9n\hat{C}_{pk;CP_i}^2} + \frac{1}{2(n-1)}} \right] \right)^{1/v}; \left(\prod_{i=1}^v \hat{C}_{pk;CP_i} \left[1 + Z_{(1+\alpha)/2} \sqrt{\frac{1}{9n\hat{C}_{pk;CP_i}^2} + \frac{1}{2(n-1)}} \right] \right)^{1/v} \right] \quad (7.2.1.e)$$

(Kushler e Hurley, 1992 [11]), com $Z_{(1+\alpha)/2}$ igual ao percentil de ordem $100(1+\alpha)/2$ % da distribuição Normal Padrão.

Contudo, conforme discussão realizada por Shinde e Khadse (2008 [15]), os limites de especificação definidos em (7.2.1.b) estão incorretos, porque assumem que os limites de especificação de CPs distintos são independentes entre si. De fato, apenas as distribuições dos CPs são independentes, porém seus limites de especificação são interrelacionados.

7.2.2 ICPM-II

Considere uma situação em que as especificações marginais das variáveis têm amplitudes $2d_i$, $i=1,2,\dots,p$. Supondo $\mathbf{X}' = (X_1, X_2, \dots, X_p) \sim N_p(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_X)$ e $\mathbf{Y}' = (Y_1, Y_2, \dots, Y_p)$, tal que os elementos de \mathbf{Y} são iguais a $Y_i = X_i/d_i$, então tem-se que \mathbf{Y} também segue distribuição normal multivariada com matriz de covariâncias $\boldsymbol{\Sigma}_Y$.

Partindo destas suposições iniciais, Veevers (1999 [22]) definiu um ICPM baseado em ACP, que é dado por:

$$MP_{PC} = \frac{1+\sqrt{2}}{6\sqrt{\lambda_1}}, \quad (7.2.2.a)$$

em que λ_1 denota o autovalor associado ao primeiro CP (ou autovetor) de $\boldsymbol{\Sigma}_Y$. Portanto, o desvio-padrão do primeiro CP de $\boldsymbol{\Sigma}_Y$ é $\sqrt{\lambda_1}$.

O índice definido na equação (7.2.2.a) mede o potencial de capacidade do processo, isto é, compara a medida de dispersão do processo ($6\sqrt{\lambda_1}$) com a quantia de dispersão permitida pelas especificações ($1+\sqrt{2}$), e é obtido tomando-se a média de outros dois índices definidos por Veevers (1999 [22]), que são dados por:

$$MP'_{PC} = \frac{1}{3\sqrt{\lambda_1}} \quad (7.2.2.b)$$

e

$$MP^*_{PC} = \frac{\sqrt{2}}{3\sqrt{\lambda_1}}. \quad (7.2.2.c)$$

O índice MP_{PC}' é construído a partir da escolha de uma hiperesfera de raio 1 para aproximar a região de especificação do processo, que é um hipercubo. Já no índice MP_{PC}'' , a hiperesfera escolhida tem raio $\sqrt{2}$. No caso, $6\sqrt{\lambda_1}$ representa o comprimento do eixo principal do elipsóide que representa a região do processo.

Em geral, é preciso estimar Σ_Y , obtendo-se, assim, estimativas de λ_1 e do próprio índice MP_{PC} .

Apesar da simplicidade e da facilidade em seu cálculo, Shinde e Khadse (2008 [15]) mostram, por meio de um interessante exemplo, que o ICPM definido na equação (7.2.2.a) pode induzir em erro os cálculos da capacidade do processo multivariado, de forma que o resultado de tal índice pode afirmar, por exemplo, que determinado processo é o mais capaz de um conjunto de processos estudados, quando na verdade é o menos capaz, ou vice-versa.

7.2.3 Método alternativo

Motivados pelas deficiências apresentadas pelos índices discutidos nas seções 7.2.1 e 7.2.2, Shinde e Khadse (2008 [15]) sugeriram um método alternativo para estimar a capacidade de processos normais multivariados, método este que se baseia na distribuição de probabilidade empírica dos CPs.

Suponha que $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ e considere as definições (notações) a seguir:

$$\mathbf{T}'_X = (T_1, T_2, \dots, T_p) : \text{vetor-alvo para } \mathbf{X}$$

$$\mathbf{LIE}'_X = (LIE_1, LIE_2, \dots, LIE_p) : \text{vetor de especificação inferior para } \mathbf{X}$$

$$\mathbf{LSE}'_X = (LSE_1, LSE_2, \dots, LSE_p) : \text{vetor de especificação superior para } \mathbf{X}$$

$$S = \{\mathbf{x} \mid \mathbf{LIE}'_X \leq \mathbf{x} \leq \mathbf{LSE}'_X\} : \text{região de especificação hiperretangular para } \mathbf{X}$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p : \text{autovalores de } \Sigma$$

$$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p : \text{autovetores de } \Sigma$$

$$Y_i = \mathbf{u}'_i \mathbf{X} : i\text{-ésimo CP, } i = 1, 2, \dots, p$$

$$\mathbf{Y} = \mathbf{U}' \mathbf{X} : \text{vetor de CPs, onde } \mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p)$$

$$E(\mathbf{Y}) = \mathbf{U}'\boldsymbol{\mu} \text{ e } Cov(\mathbf{Y}) = Diag(\lambda_1, \lambda_2, \dots, \lambda_p)$$

$$\mathbf{T}_Y = \mathbf{U}'\mathbf{T}_X : \text{vetor-alvo para } \mathbf{Y}$$

$$V = \{\mathbf{y} \mid \mathbf{LIE}_X \leq \mathbf{U}\mathbf{y} \leq \mathbf{LSE}_X\} : \text{região de especificação para } \mathbf{Y}$$

Note que V , ao contrário da região de especificação S , não é hiperretangular, e além disso, é complexa, já que se trata de um conjunto de $2p$ desigualdades lineares em p variáveis. Observe também que:

$$Y_i \sim N(\mathbf{u}'_i \boldsymbol{\mu}, \lambda_i), \quad i = 1, 2, \dots, p \text{ e } Y_1, Y_2, \dots, Y_p \text{ são independentes.}$$

Considere os primeiros ν CPs, Y_1, \dots, Y_ν , que, juntos, explicam cerca de 90% da variabilidade do processo. A fim de estudar a capacidade de processos normais multivariados com base nos primeiros ν CPs apenas, inicialmente deve-se encontrar a região de especificação V' para os primeiros ν CPs Y_1, \dots, Y_ν , fixando $Y_i = E(Y_i)$ para $i = \nu + 1, \dots, p$. Uma possível justificativa para esta última passagem é que os CPs $Y_{\nu+1}, \dots, Y_p$ apresentam variação muito menor se comparados com os primeiros ν CPs.

Então, tem-se que:

$$V' = \left\{ (y_1, y_2, \dots, y_\nu) \mid \begin{array}{l} \mathbf{LIE}_X \leq \mathbf{U}\mathbf{y} \leq \mathbf{LSE}_X, \text{ onde } \mathbf{y}' = (y_1, y_2, \dots, y_p) \\ \text{tal que } y_r = E(Y_r), \quad r = \nu + 1, \dots, p \end{array} \right\} \quad (7.2.3.a)$$

Os ICPMs baseados em probabilidades, segundo Shinde e Khadse (2008 [15]), são definidos como segue:

$$Mp_1 = P\left\{ \mathbf{Y} = (Y_1, Y_2, \dots, Y_k)' \in V' \mid \mathbf{Y} \sim N_k(\boldsymbol{\mu}_Y = \mathbf{T}_Y, \Sigma_Y = Diag(\lambda_1, \dots, \lambda_k)) \right\} \quad (7.2.3.b)$$

$$Mp_2 = P\left\{ \mathbf{Y} = (Y_1, Y_2, \dots, Y_k)' \in V' \mid \mathbf{Y} \sim N_k(\boldsymbol{\mu}_Y, \Sigma_Y = Diag(\lambda_1, \dots, \lambda_k)) \right\}$$

Note que o Mp_1 é análogo ao MC_p e o Mp_2 é análogo ao MC_{pk} . Se $Mp_1 \geq 0,9973$, o processo é dito potencialmente capaz (o potencial de capacidade refere-se à variação, mas não à locação). Logo, não há necessidade de promover a diminuição na variabilidade do processo. Se $Mp_2 \geq 0,9973$, então o processo é atualmente capaz (a performance atual de capacidade leva em consideração o alvo do processo). Portanto, não é preciso ajustar (adequar) o alvo para a média do processo, ou seja, não é necessário alterar o seu posicionamento central.

Como a região de especificação V' para (Y_1, \dots, Y_v) , definida em (7.2.3.a), é complexa e é um conjunto de $2v$ desigualdades, torna-se difícil o cálculo dos índices Mp_1 e Mp_2 , devido à dificuldade na avaliação de integrais múltiplas em regiões complicadas. Uma saída encontrada por Shinde e Khadse (2008 [15]) para este problema consiste em uma aproximação empírica, que funciona por meio dos seguintes passos:

1. Gerar duas amostras aleatórias de tamanho grande N ($N \geq 20000$) da distribuição dos primeiros k CPs, com vetor de médias como segue:

Amostra I: Vetor de médias igual ao vetor-alvo, isto é, $\mathbf{T}'_y = (T_{Y_1}, T_{Y_2}, \dots, T_{Y_k})$

Amostra II: Vetor de médias $\boldsymbol{\mu}'_y = (\mu_{Y_1}, \mu_{Y_2}, \dots, \mu_{Y_k})$

2. Estimar Mp_1 e Mp_2 por:

$$\hat{M}p_1 = \frac{\text{Número de observações } \mathbf{y} = (y_1, y_2, \dots, y_k)' \text{ da Amostra I } \ni \mathbf{y} \in V'}{N} \quad (7.2.3.c)$$

$$\hat{M}p_2 = \frac{\text{Número de observações } \mathbf{y} = (y_1, y_2, \dots, y_k)' \text{ da Amostra II } \ni \mathbf{y} \in V'}{N}$$

Embora seja uma técnica antiga e muito usada na prática, a aproximação empírica permite também que sejam consideradas outras regiões de especificação além da hiperretangular. De fato, regiões de especificação não-hiperretangulares não são consideradas em outros ICPMs existentes, uma vez que tornam a implementação computacional de tais índices uma tarefa complicada.

7.3 Exemplo 14

Considere novamente o conjunto de dados analisado no Exemplo 12 (Tabela 6.3 do Apêndice B). Elimine as 5 últimas amostras, que estão fora de controle, e desconsidere as especificações (referências) dos parâmetros do processo. Suponha que a empresa em questão deseja verificar se o processo se encontra sob controle e, em caso positivo, estudar a sua capacidade, considerando os limites de especificação dados por $\mathbf{LIE}' = (LIE_1, LIE_2, LIE_3) = (85, 85, 85)$ e $\mathbf{LSE}' = (LSE_1, LSE_2, LSE_3) = (115, 115, 115)$.

Nas Figuras 7.3.a, 7.3.b, 7.3.c, 7.3.d, 7.3.e e 7.3.f, é apresentada uma sequência de passos para a análise do conjunto de dados em questão no sistema de CEP *on-line* desenvolvido, desde a inserção dos dados (Figura 7.3.a) até a visualização do relatório instantâneo para a capacidade (Figura 7.3.f).

A Figura 7.3.d apresenta os gráficos de controle para a variabilidade e o Gráfico T^2 . Observa-se que o processo encontra-se sob controle, tanto em relação à variabilidade quanto à média, pois não há pontos que excedam o limite superior de controle em nenhum dos gráficos construídos. Portanto, pode-se estudar a capacidade desse processo. A Figura 7.3.d apresenta ainda alguns resultados da Análise de Componentes Principais que, por sua vez, auxiliam na escolha do número ideal de componentes para a análise. Tais resultados, no caso, sugerem a escolha dos dois primeiros componentes, visto que os mesmos explicam aproximadamente 82% da variabilidade do processo, porcentagem esta condizente com os propósitos da análise realizada.

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (n é constante):

Defina o nome da sua tabela de dados:

Defina o número de variáveis do processo:

Cole aqui os dados referentes a cada variável, sendo a primeira linha o nome da variável:

v1	v2	v3
V1	V2	V3
100.84	99.29	99.19
99.15	98.96	100.35
98.25	99.63	100.17
98.31	100.83	97.84
99.99	100.50	101.34
100.04	99.77	101.77
97.94	99.50	98.64
101.96	100.41	99.60
99.85	99.95	99.77
99.89	100.43	100.36
101.31	101.11	102.08
100.23	99.07	101.93

Figura 7.3.a: Inserindo os dados.

O conjunto de dados inserido no banco contém as seguintes informações:

Nome do conjunto de dados:	exemplo14
Número de variáveis:	3
Número de observações:	240
Número de amostras:	20
Tamanho das amostras:	12

Se as informações acima estão corretas, clique no botão 'Avançar' para dar continuidade à análise.

Caso haja alguma irregularidade nas informações, por favor [clique aqui](#) para inserir novamente os dados.

Avançar

Figura 7.3.b: Confirmando os dados inseridos.

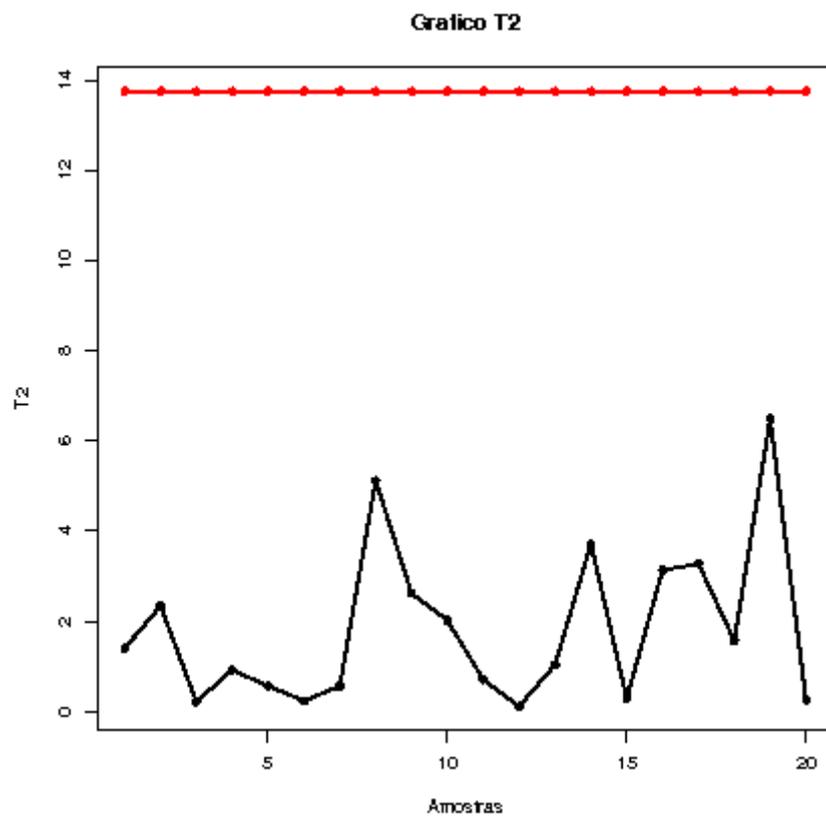
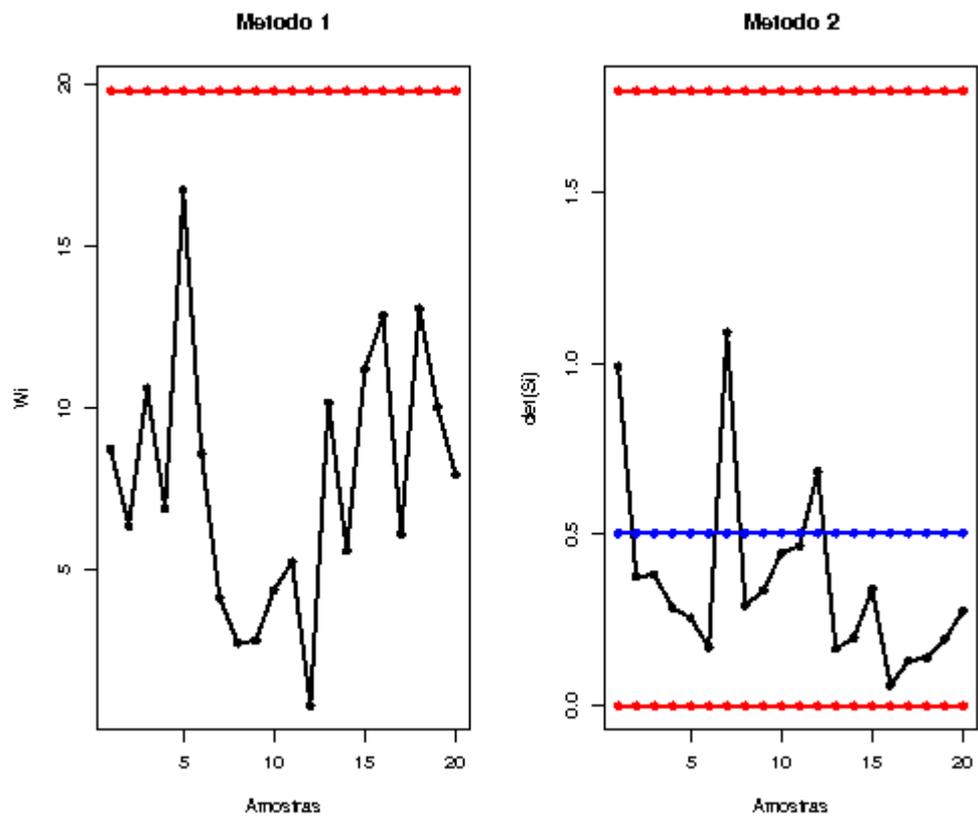
Defina as condições para os gráficos: Sem especificação

Defina a etapa atual da análise*: Etapa 1

* A Primeira Etapa (ou Etapa 1) da análise visa a identificação das amostras fora de controle pelo gráfico T2, enquanto a Segunda Etapa (ou Etapa 2) visa o monitoramento futuro do processo, quando retiradas todas as amostras fora de controle, com a obtenção de estimativas confiáveis do vetor de médias, da matriz de covariâncias e do limite superior de controle.

Avançar

Figura 7.3.c: Definindo as características e condições do processo.



Análise de Componentes Principais:

	Componente1	Componente2	Componente3
Lambda	1.8223	0.6456	0.5321
Proporção da Variância	0.6074	0.2152	0.1774
Proporção Acumulada	0.6074	0.8226	1

Figura 7.3.d: Relatório instantâneo gerado para os gráficos de controle da variabilidade e Gráfico T^2 (dados agrupados), junto com alguns resultados da Análise de Componentes Principais.

Se o processo encontra-se sob controle, defina os limites inferiores e superiores de especificação das variáveis V1, V2, V3, nessa ordem e separados por espaços simples. Defina também o nível de confiança, *alpha*, dos intervalos para cada índice, bem como o número ideal, *v*, de componentes principais a serem usados na análise da capacidade do processo.

LIE:

LSE:

Alpha*:

v^{**} :

Figura 7.3.e: Definindo as informações necessárias para o estudo da capacidade do processo.

	Limite Inferior	Estimativa Pontual	Limite Superior
MCp	2.58	2.84	3.09
MCpk	2.56	2.82	3.08
MCpm	-	2.83	-
MPpc	-	4.6	-
Mp1	-	0.9998	-
Mp2	-	0.9996	-

Figura 7.3.f: Relatório instantâneo gerado para a capacidade.

A Figura 7.3.f apresenta os índices de capacidade calculados com base nos dois primeiros componentes principais (que explicam 82% da variabilidade do processo). Observa-se que os índices MC_p , MC_{pk} , MC_{pm} e MP_{pc} assumem valores maiores do que 1, o que é indicativo de que o processo dispõe de boa capacidade para atender às especificações estabelecidas. Já os índices da aproximação empírica, Mp_1 e Mp_2 , apresentaram valores são maiores do que 0,9973. Portanto, o processo em questão é potencialmente e atualmente capaz.

7.4 Considerações Finais

Neste capítulo foram discutidos alguns dos principais índices de capacidade para dados de processos normais multivariados, que podem ser obtidos usando os componentes principais.

Não existe ainda um consenso na literatura sobre qual dos índices aqui apresentados seria o melhor ou mesmo sobre como a capacidade, em termos multivariados, deveria ser quantificada. No entanto, a Análise de Componentes Principais mostrou-se uma técnica altamente promissora para aplicações (estudos) envolvendo dados de processos multivariados, visto que é de fácil implementação e fornece meios relativamente simples de se obter índices de capacidade para processos multivariados.

É importante ressaltar que os índices aqui discutidos são melhores usados em modo comparativo, isto é, para fins de melhoria de processo, e devem sempre ser interpretados com cuidado como valores individuais.

No capítulo 8 é apresentado o sistema de CEP *on-line* desenvolvido ao longo deste projeto de mestrado. São apresentados, em detalhes, e separados por tipo de metodologia usada (CEP univariado e CEP multivariado), todos os passos envolvidos no uso do sistema, desde a inserção dos dados até a visualização dos relatórios instantâneos com os gráficos de controle e os índices de capacidade do processo.

Capítulo 8

Controle do processo *on-line*

Como já foi citado no início, o intuito deste trabalho de mestrado é estudar técnicas estatísticas que possam ser usadas no controle de processos e aplicadas num sistema *on-line* inovador, que proporciona uma análise rápida, sucinta e objetiva da situação de processos univariados e também multivariados. Este capítulo é voltado para o esclarecimento da utilização do sistema de CEP *on-line* desenvolvido, descrevendo todas as etapas envolvidas no uso do mesmo, desde a inserção dos dados até a geração dos relatórios instantâneos para os gráficos de controle e para a capacidade. Para o caso em que se considera apenas uma característica da qualidade do processo, tal sistema foi projetado tanto para a geração de gráficos para o controle de variáveis mensuráveis (Gráficos de Shewhart), quanto para gráficos voltados para o controle de atributos, além da possibilidade de se realizar também o estudo da capacidade do processo, caso o mesmo se encontre sob controle. Para o caso em que são observadas duas ou mais características da qualidade, o sistema proposto permite a geração de gráficos voltados para o controle da variabilidade do processo e também de gráficos para o controle das médias das variáveis observadas (Gráfico T^2 de Hotelling), sendo possível também analisar a capacidade do processo multivariado, por meio de alguns índices encontrados na literatura recente. Não foi encontrado na literatura nenhum outro sistema que disponibilize ao usuário ferramentas de CEP multivariado, o que as torna uma exclusividade do sistema *on-line* aqui proposto.

Além disso, os *softwares* utilizados para a realização de estudos de CEP, geralmente são pagos, inibindo, muitas vezes, a utilização das ferramentas estatísticas apresentadas neste relatório, por parte de certos tipos e tamanhos de empresas. Assim, a proposta deste trabalho é desenvolver um sistema alternativo, acessível à maioria das empresas, considerando a aplicação das ferramentas de controle da qualidade no *software* livre denominado *R* que, junto a outros *softwares* (os quais são apresentados a seguir), proporcionam resultados altamente satisfatórios, em questões de interação do usuário com o programa, se apresentando como uma plataforma amigável, de fácil utilização (principalmente por pessoas com pouco conhecimento estatístico, como é o caso de profissionais da linha de produção) e baixo custo (somente necessita de um computador conectado à Internet para o seu uso).

Os demais *softwares* livres utilizados, além do *R*, são o servidor *Apache* versão 2.2.3 e o manipulador de banco de dados, denominado *phpMyAdmin* versão 2.7.0-pl2. Também foram consideradas *web* linguagens, tais como o *PHP* versão 5.2.1 e *HTML*, usadas como base para a montagem das páginas e conexões com o banco; o *MySQL* versão 5.0.38, usado para manipular o banco de dados; e o sistema operacional *Linux* com distribuição *Ubuntu 7.04 Feisty Fawn*. Essas ferramentas são essenciais para a conexão entre a inserção de dados, análise dos mesmos e geração de relatório, sendo que a metodologia computacional aqui considerada é balizada pelo sistema desenvolvido por Souza & Louzada-Neto (2008 [16]).

Uma ilustração do procedimento de inserção de dados até a visualização do relatório instantâneo é dada na Figura 8.a, logo abaixo.

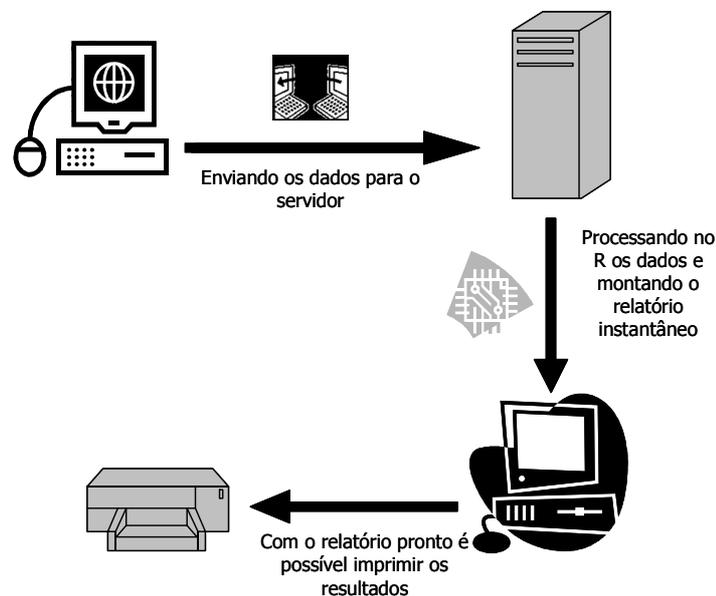


Figura 8.a: Esquema da geração do relatório instantâneo.

O sistema aqui sugerido é de livre acesso e está disposto em um endereço na Internet, mas poderia ser implantado, alternativamente, em um servidor interno de uma indústria ou estabelecimento comercial.

A seguir, é descrito o funcionamento do sistema de CEP *on-line*, separado por tipo de metodologia usada: CEP univariado (seção 8.1) ou CEP multivariado (seção 8.2).

8.1 Procedimentos *on-line* de CEP univariado

Para os casos em que deve ser utilizada a metodologia de CEP univariado, o sistema de CEP *on-line* funciona com os seguintes passos:

- 1- Inserção no banco de dados e de características necessárias para a criação da tabela no banco do servidor, tais como: número de variáveis, número de linhas dos dados e nome da tabela a ser inserida;
- 2- Confirmação das informações dadas na página anterior;
- 3- Seleção do tipo de dados que foi inserido: quantitativo (variáveis) ou qualitativo (atributos); escolha dos tipos de gráficos a serem gerados, de acordo com o tipo de dado inserido e definição se existem ou não especificações da indústria ou empresa em questão;
- 4- Visualização dos resultados para gráficos de controle e possível impressão;
- 5- Visualização dos resultados para a capacidade e possível impressão.

No tópico a seguir são detalhados os procedimentos de inserção dos dados quantitativos no sistema.

8.1.1 Entrada de dados

O procedimento de inserção dos dados tem uma sequência de particularidades a serem seguidas, de tal sorte que os dados sejam inseridos corretamente no banco e os relatórios instantâneos possam ser gerados.

Primeiramente, é preciso ter total conhecimento das amostras coletadas, isto é, ter conhecimento do número de amostras e do tamanho de cada amostra.

Ao entrar na página de inserção dos dados, representada na Figura 8.1.1.a, observam-se cinco campos de preenchimento livres que são detalhados a seguir:

1. Este campo pede o número de amostras coletadas no processo;

2. Pede o tamanho de cada amostra, caso os tamanhos das amostras sejam iguais;
3. Pede o vetor de tamanhos das amostras, caso os tamanhos das amostras sejam variáveis;
4. Neste campo é pedido o nome da tabela a ser criada no banco do servidor, tal que o nome dado não pode ser igual a nenhum que já esteja no banco de dados (caso o usuário entre com um nome existente, na página seguinte será alertado pelo sistema);
5. Espaço para digitar ou colar os dados a serem avaliados.

Página 3

Sistema de Controle Estatístico de Processo

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado: 1

Defina o tamanho de cada amostra (se n é constante): 2

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço.

3

Defina o nome da sua tabela de dados: 4

Cole aqui seus dados, sendo a primeira linha o nome da variável:

5

Avançar

Figura 8.1.1.a: Página de inserção dos dados.

Os itens 2 e 3 não podem ser preenchidos ao mesmo tempo, isto é, quando se tem o mesmo tamanho de amostra é preciso preencher o item 2 e quando se tem tamanhos variáveis de amostra é preciso preencher o item 3. Quando o tamanho das amostras é variável, um vetor com o tamanho de cada amostra deve ser digitado ou

colado no item 3 com apenas um espaço entre os números, como mostra a Figura 8.1.1.b.

OU cole aqui um vetor com o tamanho de cada amostra, onde cada amostra é separada por um espaço:

```
16 14 15 16 15 14 15 15 16
16 15 15 15 15 14 15 16 16
15 15 16 16 16 15 14
```

Figura 8.1.1.b: Exemplo de inserção do vetor de tamanhos das amostras.

Caso o nome da tabela já exista no banco de dados do servidor, a saída é voltar à página anterior (página de inserção dos dados) e modificar o nome até que algum seja aceito.

Para digitar ou colar os dados de tal forma que estes sejam inseridos corretamente, deve-se organizá-los de maneira que fique apenas uma coluna, a qual contém os valores coletados de maneira que estejam dispostos por amostra e por sub-amostra, nessa sequência, ou seja, os valores são dispostos por cada amostra, sendo repetidos caso n seja maior que 1 e de acordo com a numeração das sub-amostras, que vai de 1 a n_i dentro de cada amostra.

É importante ressaltar que, ao entrar com os dados na página de inserção (item 5), a primeira linha deve conter o nome da variável, sendo que este não pode conter espaços. Depois de preenchidos todos os campos, basta clicar em “Avançar” para ir à próxima página.

8.1.2 Confirmação das informações inseridas

O objetivo desta página é apresentar as informações referentes ao conjunto de dados inserido anteriormente, de tal sorte que se verifique se os dados foram inseridos corretamente. As informações que aparecem são tais como as apresentadas na Figura 8.1.2.a.

Caso as informações apresentadas estejam corretas, basta clicar em “Avançar” para ir ao próximo passo. Caso contrário, deve-se clicar no *link* “[clique aqui](#)” para inserir novamente os dados.

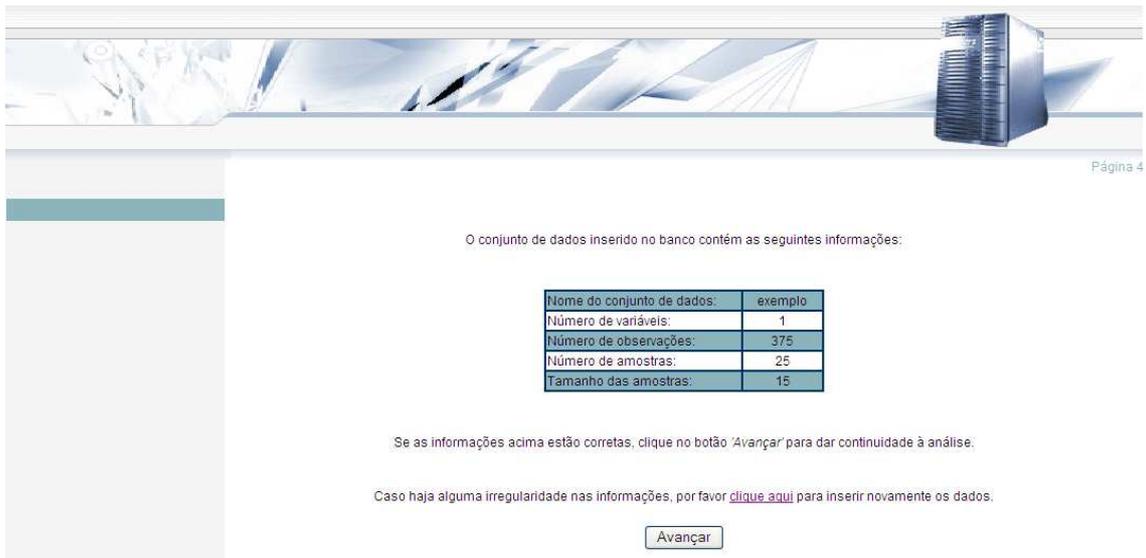


Figura 8.1.2.a: Exemplo de confirmação de informações.

8.1.3 Seleção das características para a geração dos relatórios instantâneos

Nesta página, o usuário deve selecionar primeiramente o tipo de conjunto de informações monitoradas, isto é, se os dados inseridos são de natureza quantitativa (*Variáveis*), seguindo para a realização de gráficos de controle para variáveis, ou de natureza qualitativa (*Atributos*), seguindo então para a realização dos gráficos de controle para atributos, tal como na Figura 8.1.3.a.

Após essa seleção, o usuário deve informar, de acordo com o tipo de dados inseridos, os gráficos que devem ser gerados. Caso os dados sejam quantitativos (*Variáveis*), os gráficos que podem ser feitos são o S com o \bar{X} (*Gráficos S e Xbarra*) ou o R com o \bar{X} (*Gráficos R e Xbarra*). Caso os dados sejam qualitativos (*Atributos*), os gráficos podem ser o p (*Gráfico p*), o np (*Gráfico np*), o c (*Gráfico c*) ou o u (*Gráfico u*).

O passo seguinte para a seleção de características é a definição das condições para os gráficos, isto é, definir se existem (*Com especificação*) ou não (*Sem especificação*) especificações com relação aos parâmetros. Caso existam especificações, é necessário preencher os campos em branco que aparecem (*Média e Desvio-Padrão*, caso os dados inseridos sejam do tipo *Variáveis*; *Proporção*, caso o gráfico a ser construído seja o *Gráfico p* ou o *Gráfico np*; *Número de Defeitos*, se for selecionado o *Gráfico c*; e *Número Médio de Defeitos*, se for escolhido o *Gráfico u*).

Finalmente, basta clicar em “Avançar”, para então gerar o relatório instantâneo da situação do processo.

Escolha o tipo de dados que você inseriu: Variáveis

Escolha os gráficos adequados para seu tipo de dados: Gráficos S e Xbarra

Defina as condições para cada gráfico: Sem especificação

Avançar

Figura 8.1.3.a: Exemplo de seleção de características.

8.1.4 Relatório instantâneo para gráficos de controle

O relatório instantâneo gerado contém as figuras dos gráficos de controle construídos para os dados inseridos e algumas informações adicionais, como por exemplo, a data e o horário da geração deste relatório. Uma ilustração desse relatório instantâneo é dada na Figura 8.1.4.a.

8.1.5 Relatório instantâneo para a capacidade

Para o caso em que os dados inseridos são do tipo *Variáveis* e o processo em questão se encontra sob controle, há ainda outro relatório que pode ser gerado. Trata-se do relatório instantâneo da capacidade do processo. Para a sua geração, é necessário que na página que contém o relatório instantâneo para gráficos de controle, sejam inseridos no formulário à esquerda (como mostra a Figura 8.1.4.a) os valores dos limites inferior (*LIE*) e superior (*LSE*) de especificação, bem como o valor do nível de confiança (*Alpha*), em número decimal e com “.” para separação das casas, dos intervalos a serem construídos. O relatório gerado mostra um gráfico histograma para a verificação de centralidade dos dados e, abaixo dele, as estimativas pontuais dos índices C_p , C_{pk} , C_{pm} , P e *Desvio* (que é a diferença máxima entre a média e o valor-alvo do processo), bem

como os intervalos ao nível de 100(α) % de confiança para os índices C_p , C_{pk} e P . Uma ilustração desse relatório instantâneo é dada na Figura 8.1.5.a.

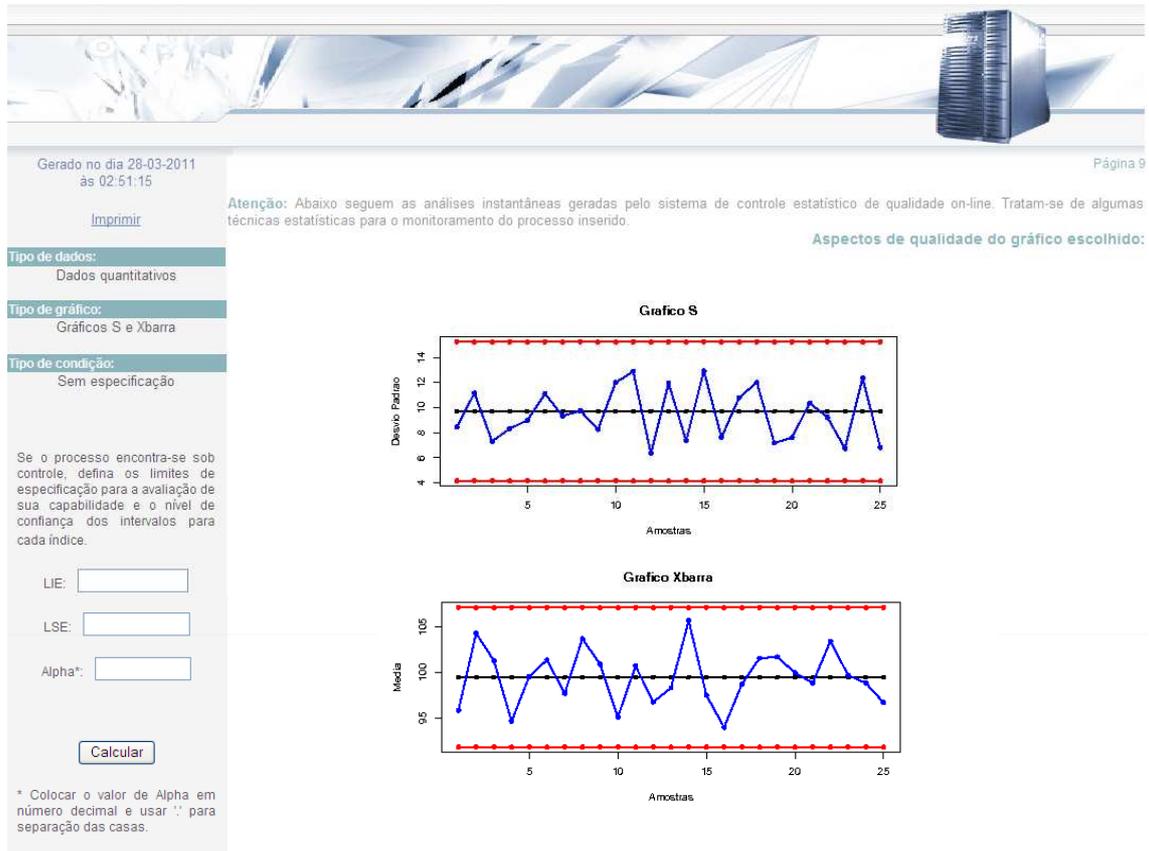


Figura 8.1.4.a: Ilustração do relatório instantâneo para gráficos de controle.

8.2 Procedimentos *on-line* de CEP multivariado

Para os casos em que deve ser utilizada a metodologia de CEP multivariado, o sistema de CEP *on-line* funciona com os seguintes passos:

- 1- Inserção no banco de dados e de características necessárias para a criação da tabela no banco do servidor, tais como: número de variáveis, número de linhas dos dados e nome da tabela a ser inserida;
- 2- Confirmação das informações dadas anteriormente;

- 3- Definição se existem ou não especificações da indústria ou empresa em questão e definição também do objetivo (etapa) da análise, caso tais especificações não existam;
- 4- Visualização dos resultados para gráficos de controle e possível impressão;
- 5- Visualização dos resultados para a capacidade e possível impressão.

No tópico a seguir são detalhados os procedimentos de inserção dos dados quantitativos no sistema.

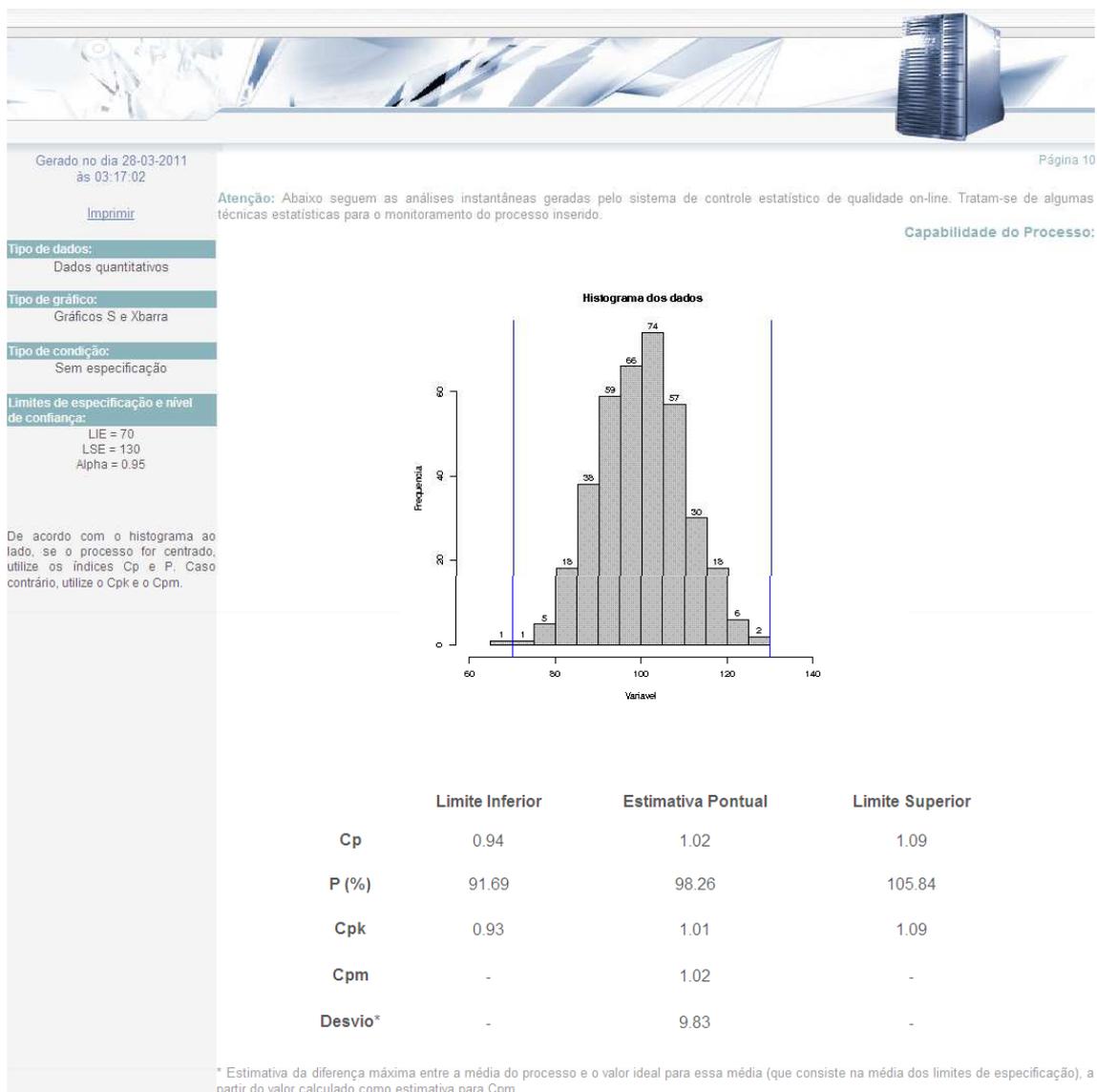


Figura 8.1.5.a: Ilustração do relatório instantâneo para a capacidade.

8.2.1 Entrada de dados

Assim como para o caso univariado (seção 8.1.1), o procedimento de inserção dos dados multivariados tem uma sequência de particularidades a serem seguidas, de modo que os dados sejam inseridos corretamente no banco e os relatórios instantâneos possam ser gerados

Primeiramente, é preciso ter total conhecimento das amostras coletadas, isto é, ter conhecimento do número de amostras e do tamanho de cada amostra.

Ao entrar na página de inserção dos dados, representada na Figura 8.2.1.a, observam-se quatro campos de preenchimento livres, que são detalhados a seguir:

1. Este campo pede o número de amostras coletadas no processo;
2. Pede o tamanho de cada amostra, tal que os tamanhos das amostras devem ser iguais;
3. Neste campo é pedido o nome da tabela a ser criada no banco do servidor, tal que o nome dado não pode ser igual a nenhum que já esteja no banco de dados (caso o usuário entre com um nome existente, na página seguinte será alertado pelo sistema);
4. Este campo pede o número de variáveis do processo.

Ao preencher o campo referente ao item 4, surgem, logo em seguida, novos campos abaixo dele, como mostra a Figura 8.2.1.b, nos quais devem ser digitados ou colados os valores de cada variável, de forma que tais valores formem uma única coluna dentro de cada um desses campos, e a primeira linha de cada campo contenha o nome da variável em questão, o qual não pode conter espaços.

Caso o nome da tabela já exista no banco de dados do servidor, a saída é voltar à página inicial de inserção dos dados (que é aquela em que os campos para inserção dos valores das variáveis em estudo ainda não aparecem), e modificar o nome até que algum seja aceito.

Assim, preenchidos os quatro campos iniciais e inseridos os valores de cada uma das variáveis do processo, basta clicar em “Avançar” para ir ao próximo passo.

Página 3

Sistema de Controle Estatístico de Processo

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (n é constante):

Defina o nome da sua tabela de dados:

Defina o número de variáveis do processo:

Figura 8.2.1.a: Página de inserção dos dados, antes de selecionar o número de variáveis do processo.

Página 4

Sistema de Controle Estatístico de Processo

Por favor, preencha os campos abaixo para prosseguir com a análise.

Defina o número de amostras utilizado:

Defina o tamanho de cada amostra (n é constante):

Defina o nome da sua tabela de dados:

Defina o número de variáveis do processo:

Cole aqui os dados referentes a cada variável, sendo a primeira linha o nome da variável:

v1

v2

v3

Figura 8.2.1.b: Exemplo de página de inserção dos dados quando o número de variáveis do processo é igual a 3.

8.2.2 Confirmação das informações inseridas

Nesta página são apresentadas as informações referentes ao conjunto de dados inserido anteriormente, de modo que se verifique se os dados foram inseridos corretamente. As informações que aparecem são tais como as mostradas na Figura 8.2.2.a.

Caso as informações apresentadas estejam corretas, basta clicar em “Avançar” para ir ao próximo passo. Caso contrário, deve-se clicar no *link* “[clique aqui](#)” para inserir novamente os dados.

Nome do conjunto de dados:	exemplo
Número de variáveis:	3
Número de observações:	240
Número de amostras:	20
Tamanho das amostras:	12

Figura 8.2.2.a: Exemplo de confirmação das informações inseridas.

8.2.3 Seleção das características para a geração dos relatórios instantâneos

Nesta página, como ilustra a Figura 8.2.3, o usuário deve definir as condições para os gráficos a serem construídos, isto é, definir se existem (*Com especificação*) ou não (*Sem especificação*) especificações com relação aos parâmetros do processo. Caso existam especificações, é necessário preencher os campos em branco que surgem, referentes à média de cada uma das variáveis observadas e às variâncias e covariâncias entre elas. Caso não existam especificações, o usuário deve definir a etapa atual da análise, de acordo com os seus interesses: *Etapa 1*, se o objetivo for identificar as amostras que estão sob controle, ou *Etapa 2*, caso o processo esteja sob controle e haja

interesse em usar tais amostras para estimar o vetor de médias, a matriz de covariâncias e o limite superior de controle a serem usados no monitoramento futuro do processo.

Em seguida, basta clicar em “Avançar”, para então gerar o relatório instantâneo da situação do processo.



Figura 8.2.3.a: Exemplo de seleção das características.

8.2.4 Relatório instantâneo para gráficos de controle

O relatório instantâneo gerado contém as figuras dos gráficos de controle para a variabilidade, construídos segundo os dois procedimentos de controle da variabilidade discutidos na seção 6.1, mais o Gráfico T^2 . Para o caso em que as amostras coletadas são de tamanhos iguais a um, os gráficos para a variabilidade não são exibidos e são mostrados apenas os Gráficos T^2 construídos a partir de dois estimadores distintos da matriz de covariâncias discutidos na seção 6.2.2. Logo abaixo dos gráficos, é apresentada uma tabela contendo alguns resultados da Análise de Componentes Principais, como a variância (λ) de cada componente, as proporções e as proporções acumuladas da variabilidade total (ou original) do processo que são explicadas por cada um deles. Tais resultados auxiliam na escolha do número ideal de componentes a serem usados na análise da capacidade do processo. Caso no passo (página) anterior fosse selecionada a opção *Etapa 2*, seriam aqui exibidos, adicionalmente, os valores das estimativas dos parâmetros e também do limite superior de controle a ser usado no monitoramento futuro do processo.

Uma ilustração desse relatório instantâneo é dada na Figura 8.2.4.a.

8.2.5 Relatório instantâneo para a capacidade

Caso o processo em questão esteja sob controle estatístico, há ainda outro relatório importante que pode ser gerado. Trata-se do relatório instantâneo da capacidade do processo. Para a sua geração, é necessário que na página anterior (aquela que contém o relatório instantâneo para gráficos de controle) sejam inseridos no formulário à esquerda (como mostra a Figura 8.2.4.a) os vetores de limites inferiores (*LIE*) e superiores (*LSE*) de especificação, tal que os valores para cada um deles estejam separados por espaços simples. Devem ser inseridos também, neste formulário, o nível de confiança (*Alpha*) dos intervalos a serem construídos (em número decimal e com “.” para a separação das casas) e o número ideal de componentes principais (*v*) a serem usados na análise da capacidade do processo. O relatório gerado mostra as estimativas pontuais dos índices MC_p , MC_{pk} , MC_{pm} , MP_{pc} , Mp_1 e Mp_2 , bem como os intervalos ao nível de $100(Alpha)\%$ de confiança para os índices MC_p e MC_{pk} . Uma ilustração desse relatório instantâneo é dada na Figura 8.2.5.a.



Gerado no dia 28-03-2011
às 05:50:43

Página 8

[Imprimir](#)

Atenção: Abaixo seguem as análises instantâneas geradas pelo sistema de controle estatístico de qualidade on-line. Tratam-se de algumas técnicas estatísticas para o monitoramento do processo multivariado insendo.

Aspectos de qualidade dos gráficos multivariados:

Tipo de dados:
Dados quantitativos multivariados

Tipo de gráfico:
Gráficos de controle da variabilidade e Gráfico T2 de Hotelling

Tipo de condição:
Sem especificação (etapa 1)

Se o processo encontra-se sob controle, defina os limites inferiores e superiores de especificação das variáveis V1, V2, V3, nessa ordem e separados por espaços simples. Defina também o nível de confiança, alpha, dos intervalos para cada índice, bem como o número ideal, v, de componentes principais a serem usados na análise da capacidade do processo.

LIE:

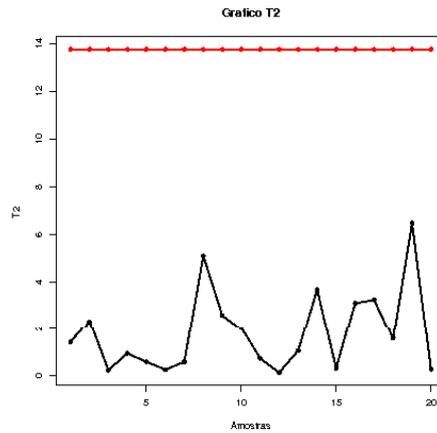
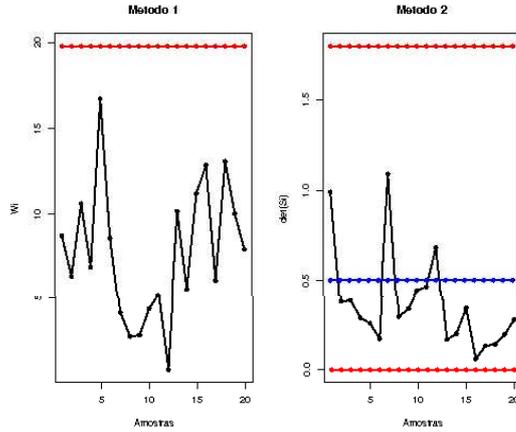
LSE:

Alpha*:

v**:

* Colocar o valor de Alpha em número decimal e usar '.' para separação das casas.

** Defina v com base na Análise de Componentes Principais ao lado, sendo v o número de componentes que explicam por volta de 90% da variabilidade original dos dados.



Análise de Componentes Principais:

	Componente1	Componente2	Componente3
Lambda	1.8223	0.6456	0.5321
Proporção da Variância	0.6074	0.2152	0.1774
Proporção Acumulada	0.6074	0.8226	1

Figura 8.2.4.a: Ilustração do relatório instantâneo para gráficos de controle.



Figura 8.2.5.a: Ilustração do relatório instantâneo para a capacidade.

Capítulo 9

Conclusões e propostas futuras

Nos capítulos 2, 3 e 4 foram apresentadas duas formas de se monitorar a qualidade de um processo, seja numa indústria, num estabelecimento comercial ou em qualquer outro lugar onde o CEP seja aplicável. Estas duas formas de controle correspondem aos gráficos de controle para variáveis e gráficos de controle para atributos. Os gráficos citados são capazes de identificar irregularidades ocorridas em determinados períodos no processo estudado, de acordo com o valor de referência e os limites de controle calculados. Estas irregularidades podem ser identificadas caso os gráficos, de acordo com cada tipo, apresentem tendências, sazonalidades, pontos fora dos limites ou vários pontos consecutivos acima ou abaixo da linha de referência. A partir desses gráficos, é possível obter limites de controle que podem ser utilizados na monitoração do processo atual. Os gráficos discutidos nos capítulos 2, 3 e 4 podem ser gerados a partir de estimação ou de valores especificados (pela empresa em questão ou pelo cliente) para os parâmetros.

Além dos gráficos de controle, também foram aplicados no capítulo 5 os índices de capacidade do processo, os quais podem ser utilizados para verificar se o processo opera dentro dos limites de variação e se o mesmo é capaz de atender às exigências estabelecidas pelos clientes ou impostas pela própria empresa em questão. Tais índices são construídos baseados em algumas suposições, como a de que a variável observada siga uma distribuição Normal. Por isso, antes de se realizar um estudo de CEP, recomenda-se a verificação desta condição por meio de testes para normalidade, como Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling e Cramér-von-Mises, dentre outros. Em caso de não normalidade dos dados, recomenda-se ainda transformá-los, de modo a obter uma variável cuja distribuição possa ser bem aproximada pela Normal e, assim, poder usar os procedimentos discutidos nos capítulos 2 e 5.

No capítulo 6 foram apresentados procedimentos multivariados de CEP. De fato, em contextos reais, é natural que a qualidade de um processo seja refletida através de um conjunto de variáveis que normalmente exibem associação entre si. Nestes casos, a abordagem univariada da maioria das técnicas usualmente empregadas corresponde a uma limitação destas mesmas, o que resulta em conclusões menos precisas e coerentes. Em consequência disso, pode haver diminuição na satisfação e confiança do cliente,

aumento nas deficiências e nos custos do processo e quedas nas receitas. Foram apresentados, neste capítulo, gráficos de controle para a variabilidade, construídos segundo dois procedimentos de controle da variabilidade discutidos, e também o gráfico para controle das médias do processo, também conhecido como Gráfico T^2 de Hotelling. Tais gráficos, assim como os índices de capacidade apresentados no capítulo 7, supõem a normalidade multivariada das variáveis do processo, condição esta que deve ser verificada antes de se aplicar os procedimentos apresentados, através de testes, como por exemplo, o teste de Shapiro-Wilk generalizado para normalidade multivariada.

O uso do CEP é muito comum no meio daqueles que desejam avaliar a qualidade de seus produtos ou serviços prestados. Por este motivo, a aplicação desse estudo foi executada num sistema *on-line* com suporte no *software* estatístico *R* versão 2.5.1. Estes programas representam os cálculos dos diversos tipos de gráficos e índices discutidos neste trabalho e podem ser executados a partir de uma página na Internet ou em um servidor interno. A utilização do sistema de CEP *on-line* desenvolvido é extremamente fácil, rápida e de baixo custo (somente necessita de um computador conectado à Internet), o que o torna uma excelente solução para empresas que não têm condições de pagar por *softwares* caros que realizam análises sobre CEP. O sistema também proporciona uma rápida geração de relatórios sem ter que programar a todo o momento a realização dos gráficos e dos índices em um *software* livre.

O sistema *on-line* para CEP é de livre acesso e está disponível no endereço: 200.136.219.122/dados/tg/tg_logar.php com o *login* igual a “usuario” e a senha “tg22009”. Seu diferencial em relação a outros sistemas *on-line*, bem como para outros *softwares* estatísticos, está na disponibilidade de ferramentas para o CEP multivariado, ferramentas estas que, não até o momento, não foram encontradas em nenhum outro programa.

Como proposta de trabalho futuro, fica a implementação de gráficos para a detecção de mudanças mais sutis na média do processo, como o Gráfico de Somas Acumuladas (Cusum) e o Gráfico de Médias Móveis Exponencialmente Ponderadas (MMEP), bem como suas versões multivariadas. Pretende-se também implementar as regras de controle da Western Electric.

Referências bibliográficas

- [1] ALT, F. B. *Multivariate Quality Control*. Encyclopedia of Statistical Sciences, Vol. 6, Wiley, New York, 1985.
- [2] BARTMANN, F. C. *Idéias básicas do controle moderno de qualidade*. Campinas: UNICAMP, 78 p., 1986.
- [3] BOYLES, R. A. *The Taguchi Capability Index*. Journal of Quality Technology, Vol. 27, 1991.
- [4] BURR, I. W. *Statistical quality control methods*. New York: Marcel Dekker, 522 p., 1976.
- [5] CHAN, L. K.; CHENG, S. W.; SPIRING, F. A. *A Multivariate Measure of Process Capability*. Journal of Modeling and Simulation, 11, pg. 1-6, 1991.
- [6] CHEN, H. *A multivariate process capability index over a rectangular solid tolerance zone*. Statistica Sinica, v. 4, n. 2, p. 749-758, 1994.
- [7] DUNCAN, A. J. *Quality control and industrial statistics*. 5 ed. Homewood: Richard D. Irwin, 1047 p., 1986.
- [8] HOLMES, D. S.; MERGEN, A. E. *Improving the Performance of the T^2 Control Chart*. Quality Engineering, Vol. 5, 1993.
- [9] KANE, V. E. *Process capability indices*. Journal of Quality Technology, Vol.18, pp. 41-52, 1986.
- [10] KOTZ, S.; JOHNSON, N. L. *Process Capability Indices*. Chapman and Hall, London, 1993.

- [11] KUSHLER, R. H.; HURLEY, P. *Confidence bounds for capability indices*. Journal of Quality Technology, 24:188-95, 1992.
- [12] MONTGOMERY, D. C. *Introdução ao Controle Estatístico da Qualidade*. 4ª ed. Rio de Janeiro: Editora LTC, Tradução de Ana Maria Lima de Farias, Vera Regina Lima de Farias e Flores, Revisão Técnica de Luiz da Costa Laurencel, 513 p., 2004.
- [13] PEARN, W. L.; KOTZ, S.; JOHNSON, C. C. *Distributions and inferential properties of process capability indices*. Journal of Quality Technology, 24(4) 216-33, 1992.
- [14] SHAHRIARI, H.; HUBELE, N. F.; LAWRENCE, F. P. *A Multivariate Process Capability Vector*. Proceedings of the 4th Industrial Research Conference, Institute of Industrial Engineers, pg. 304-309, 1995.
- [15] SHINDE, R. L.; KHADSE, K. G. *Multivariate Process Capability Using Principal Component Analysis*. Wiley InterScience, 2008.
- [16] SOUZA, A. L.; LOUZADA-NETO, F. *Procedimentos Estatísticos Para Auto-Avaliação Institucional: O Caso do Departamento de Estatística da Ufscar*. Iniciação científica CNPq-PIBIC. Desenvolvida de agosto de 2006 a agosto de 2008.
- [17] SULLIVAN, J. H.; WOODALL, W. H. *A Comparison of Multivariate Quality Control Charts for Individual Observations*. Journal of Quality Technology, Vol. 27, 1995.
- [18] TAAN, W.; SUBBAIAH, P.; LIDDY, J. W. *A Note on Multivariate Capability Indices*. Journal of Applied Statistics, Vol. 20, p. 339-351, 1993.
- [19] TOLEDO, J. C.. *Qualidade Industrial: Conceitos, Sistemas e Estratégias*. São Paulo: Atlas, 1987.

- [20] TONG, L. *The Multivariate Normal Distribution*. New York: Springer-Verlag, 1990.
- [21] TRACY, N. D.; YOUNG, J. C.; MASON, R. L. *Multivariate Control Charts for Individual Observations*. Journal of Quality Technology, Vol. 24, 1992.
- [22] VEEVERS, A. *Capability Indices for Multiresponse Processes*. Statistical Process Monitoring and Optimization, Park SH, Vining GG (eds.). Marcel Dekker: New York, NY, 241-256, 1999.
- [23] VILLASENOR-ALVA, J. A.; GONZALEZ-ESTRADA, E. *A generalization of Shapiro-Wilk's test for multivariate normality*. Communications in Statistics: Theory and Methods, 38 11,1870-1883, 2009.
- [24] WANG, F. K.; CHEN, J. C. *Capability Index Using Principal Component Analysis*. Quality Engineering, 11: 21-27, 1998.
- [25] WANG, F. K.; DU, T. C. T. *Using principal component analysis in process performance for multivariate data*. Omega, 28: 185-194, 2000.
- [26] WERKEMA, M. C. C.. *Ferramentas Estatísticas Básicas para o Gerenciamento de Processos*. 1ª ed. Belo Horizonte, MG: Fundação Christiano Ottoni, Escola de Engenharia da UFMG, 404 p., 1995.
- [27] WIERDA, S. J. *A multivariate process capability index*. ASQC Quality Congress Transactions, Boston, pp. 343–348, 1993.
- [28] ZHANG, N. Z. *Estimating Process Capability Indexes for Autocorrelated Data*. Journal of Applied Statistics, Vol. 25, n.4, p. 559-574, 1998.
- [29] <http://www.pucrs.br/famat/statweb/historia/daestatistica/biografias/Shewhart.htm>

Apêndice A

Neste apêndice são apresentados os códigos de geração dos conjuntos de dados dos exemplos estudados.

Quadro 1.2: Códigos para geração dos conjuntos de dados dos exemplos discutidos.

Exemplo 1

```
m=25; n=15; mn=m*n; mi=100; sigma=10; set.seed(4095); x=round(rnorm(mn,mi,sigma),2);  
x[286:300]=x[286:300]+2*sigma
```

Exemplo 2

```
m=25; set.seed(35); n=sample(seq(14,16),m,replace=T); mn=sum(n)  
mi=100; sigma=10; set.seed(3172); x=round(rnorm(mn,mi,sigma),2)  
x[289:305]=x[289:305]+2*sigma
```

Exemplo 3

```
m=25; n=4; mn=m*n; mi=100; sigma=10  
set.seed(9962); x=round(rnorm(mn,mi,sigma),2) x[77:80]=x[77:80]+2*sigma
```

Exemplo 4

```
m1=25; mi1=100; sigma1=10; up=2  
set.seed(1769); x=round(rnorm(m1,mi1,sigma1),2)  
set.seed(8294); x[21:25]=round(rnorm(5,mi1+up*sigma1,sigma1),2)
```

Exemplo 5

```
set.seed(7256); x=rbinom(25,100,0.1); set.seed(7865)  
x[21:25]=rbinom(5,100,0.2)
```

Exemplo 6

```
tamanho=seq(90,140); set.seed(1459) tamanho=sample(tamanho,25,replace=T)  
sementes=seq(0,9999); set.seed(8667) sementes=sample(sementes,25,replace=F)  
prob=c(rep(0.1,20),rep(0.15,5)); x=numeric(25)  
for (i in 1:25) {  
  set.seed(sementes[i]); x[i]=rbinom(1,tamanho[i],prob[i])  
}
```

Exemplo 8

```
set.seed(6847); x=rpois(25,10); set.seed(1975); x[21:25]=rpois(5,20)
```

Exemplo 9

```
set.seed(4094); x=rpois(125,10); set.seed(4970) x[101:125]=rpois(25,15)
```

Exemplo 10

```
tamanho=seq(4,7); set.seed(5490); n=sample(tamanho,25,replace=T)
sementes=seq(0,9999); set.seed(3913) sementes=sample(sementes,25,replace=F)
prob=c(rep(10,20),rep(15,5)); x=0
for (i in 1:25) {
  set.seed(sementes[i]); x=c(x,rpois(n[i],prob[i]))
}
x=x[-1]
```

Exemplo 11

```
m=25; n=15; mn=m*n; mi=100; sigma=10
set.seed(55)
x=round(rnorm(mn,mi,sigma),2)
```

Exemplo 12

```
require(MASS); media1=rep(100,5); x=matrix(NA,300,5); media2=media1 media2[3]=102;
sigma=matrix(rep(0.5,25),5,5)
for (i in 1:5) sigma[i,i]=1
set.seed(5580); x[1:240,]=mvrnorm(240,media1,sigma)
set.seed(2947); x[241:300,]=mvrnorm(60,media2,sigma); x=round(x,2)
x=x[,1:3]
```

Exemplo 13

```
require(MASS); media1=rep(100,5); x=matrix(NA,25,5); media2=media1 media2[3]=103;
sigma=matrix(rep(0.5,25),5,5); for (i in 1:5) sigma[i,i]=1
set.seed(5782); x[1:20,]=mvrnorm(20,media1,sigma)
set.seed(9796); x[21:25,]=mvrnorm(5,media2,sigma); x=round(x,2)
x=x[,1:3]
```

Apêndice B

Neste apêndice estão expostos os conjuntos de dados gerados a partir dos comandos do Apêndice A.

Tabela 2.4: Valores gerados para o conjunto de dados do Exemplo 1.

Amostra	Obs	Capac.	Amostra	Obs	Capac.	Amostra	Obs	Capac.
1	1	93.26	9	6	98.67	17	11	88.47
1	2	104.76	9	7	98.52	17	12	84.77
1	3	100.82	9	8	85.29	17	13	94.20
1	4	105.82	9	9	90.96	17	14	106.47
1	5	104.63	9	10	106.45	17	15	89.05
1	6	112.62	9	11	90.24	18	1	97.67
1	7	105.47	9	12	97.33	18	2	84.54
1	8	90.99	9	13	111.30	18	3	107.27
1	9	92.10	9	14	113.33	18	4	85.95
1	10	86.53	9	15	90.77	18	5	104.28
1	11	93.65	10	1	117.23	18	6	101.39
1	12	105.68	10	2	113.99	18	7	89.88
1	13	94.46	10	3	101.19	18	8	99.64
1	14	91.57	10	4	114.51	18	9	99.43
1	15	114.12	10	5	98.30	18	10	107.01
2	1	67.99	10	6	103.41	18	11	95.66
2	2	89.62	10	7	120.91	18	12	84.19
2	3	91.06	10	8	102.96	18	13	90.69
2	4	87.67	10	9	116.19	18	14	101.27
2	5	102.40	10	10	103.59	18	15	98.16
2	6	111.66	10	11	98.63	19	1	108.26
2	7	96.73	10	12	111.04	19	2	99.13

2	8	101.51	10	13	103.63	19	3	98.91
2	9	98.92	10	14	104.56	19	4	111.48
2	10	111.38	10	15	97.39	19	5	104.05
2	11	101.65	11	1	87.27	19	6	101.70
2	12	84.28	11	2	104.93	19	7	95.89
2	13	123.24	11	3	86.12	19	8	90.69
2	14	102.24	11	4	96.21	19	9	87.89
2	15	95.91	11	5	108.78	19	10	110.86
3	1	91.13	11	6	85.17	19	11	112.76
3	2	110.86	11	7	96.41	19	12	97.62
3	3	104.80	11	8	109.12	19	13	112.36
3	4	87.93	11	9	92.76	19	14	109.96
3	5	111.69	11	10	87.30	19	15	93.48
3	6	103.90	11	11	88.67	20	1	127.62
3	7	106.21	11	12	94.48	20	2	117.85
3	8	106.07	11	13	79.93	20	3	120.94
3	9	105.41	11	14	73.96	20	4	117.69
3	10	85.17	11	15	97.29	20	5	132.35
3	11	94.54	12	1	107.72	20	6	118.84
3	12	115.25	12	2	85.64	20	7	120.85
3	13	97.46	12	3	97.12	20	8	109.00
3	14	92.61	12	4	91.18	20	9	120.42
3	15	119.21	12	5	100.22	20	10	118.72
4	1	108.33	12	6	109.32	20	11	145.68
4	2	90.37	12	7	104.35	20	12	128.44
4	3	96.91	12	8	94.81	20	13	119.59
4	4	90.64	12	9	90.55	20	14	119.39
4	5	113.32	12	10	98.16	20	15	131.14
4	6	99.12	12	11	95.35	21	1	82.51
4	7	97.99	12	12	101.16	21	2	85.90

4	8	98.14	12	13	94.44	21	3	84.72
4	9	103.02	12	14	112.12	21	4	81.13
4	10	97.39	12	15	98.81	21	5	104.65
4	11	86.77	13	1	83.60	21	6	86.10
4	12	124.66	13	2	105.12	21	7	106.91
4	13	112.02	13	3	98.46	21	8	94.42
4	14	111.40	13	4	94.72	21	9	84.35
4	15	105.61	13	5	109.93	21	10	117.64
5	1	94.07	13	6	116.19	21	11	97.28
5	2	113.84	13	7	99.91	21	12	100.59
5	3	89.17	13	8	80.54	21	13	93.01
5	4	88.42	13	9	93.09	21	14	97.87
5	5	78.49	13	10	105.72	21	15	92.72
5	6	78.96	13	11	79.25	22	1	104.50
5	7	104.79	13	12	89.16	22	2	109.41
5	8	81.31	13	13	84.02	22	3	93.29
5	9	89.94	13	14	101.51	22	4	98.24
5	10	120.46	13	15	100.27	22	5	107.84
5	11	74.66	14	1	109.57	22	6	111.92
5	12	91.82	14	2	106.60	22	7	104.90
5	13	89.30	14	3	90.37	22	8	118.31
5	14	101.84	14	4	108.59	22	9	95.33
5	15	110.87	14	5	88.72	22	10	97.57
6	1	101.05	14	6	94.51	22	11	98.45
6	2	108.56	14	7	116.26	22	12	100.04
6	3	90.89	14	8	97.05	22	13	113.75
6	4	88.89	14	9	108.87	22	14	91.66
6	5	92.72	14	10	101.07	22	15	77.17
6	6	95.57	14	11	94.51	23	1	107.92
6	7	104.71	14	12	77.17	23	2	109.43

6	8	105.00	14	13	97.84	23	3	101.25
6	9	110.99	14	14	108.35	23	4	96.21
6	10	117.08	14	15	90.55	23	5	106.25
6	11	111.73	15	1	106.13	23	6	96.31
6	12	100.69	15	2	80.95	23	7	83.09
6	13	81.43	15	3	92.08	23	8	110.67
6	14	95.72	15	4	90.02	23	9	109.84
6	15	104.75	15	5	94.36	23	10	98.67
7	1	90.00	15	6	103.58	23	11	91.59
7	2	98.84	15	7	113.44	23	12	94.85
7	3	104.63	15	8	98.02	23	13	101.74
7	4	109.49	15	9	113.03	23	14	118.98
7	5	95.81	15	10	87.80	23	15	110.83
7	6	96.25	15	11	100.72	24	1	94.60
7	7	98.98	15	12	104.07	24	2	106.53
7	8	111.73	15	13	104.65	24	3	103.27
7	9	95.46	15	14	104.60	24	4	106.43
7	10	119.46	15	15	95.58	24	5	99.88
7	11	105.64	16	1	105.94	24	6	97.18
7	12	93.25	16	2	102.36	24	7	96.01
7	13	100.26	16	3	96.74	24	8	93.65
7	14	92.57	16	4	104.16	24	9	115.23
7	15	90.42	16	5	114.92	24	10	98.06
8	1	107.42	16	6	100.49	24	11	104.72
8	2	74.22	16	7	117.11	24	12	98.94
8	3	101.55	16	8	107.43	24	13	115.69
8	4	84.11	16	9	109.78	24	14	106.67
8	5	98.80	16	10	95.75	24	15	110.59
8	6	99.62	16	11	95.16	25	1	100.50
8	7	108.77	16	12	105.90	25	2	98.65

8	8	101.08	16	13	112.03	25	3	99.54
8	9	89.03	16	14	101.46	25	4	104.91
8	10	86.85	16	15	95.78	25	5	91.70
8	11	116.28	17	1	96.51	25	6	94.68
8	12	108.27	17	2	109.58	25	7	105.16
8	13	98.76	17	3	118.75	25	8	107.26
8	14	113.51	17	4	115.63	25	9	105.08
8	15	98.00	17	5	96.48	25	10	107.73
9	1	98.47	17	6	111.28	25	11	95.04
9	2	101.70	17	7	102.37	25	12	100.09
9	3	94.52	17	8	99.89	25	13	100.60
9	4	112.74	17	9	95.27	25	14	102.59
9	5	88.21	17	10	114.00	25	15	113.99

Tabela 2.5: Valores gerados para o conjunto de dados do Exemplo 2.

Amostra	Obs	Capac.	Amostra	Obs	Capac.	Amostra	Obs	Capac.
1	1	104.26	9	8	99.80	17	14	93.05
1	2	102.09	9	9	105.78	17	15	91.00
1	3	103.13	9	10	103.91	17	16	106.67
1	4	100.08	9	11	105.28	18	1	95.35
1	5	105.61	9	12	86.40	18	2	110.69
1	6	114.14	9	13	90.62	18	3	112.62
1	7	93.78	9	14	101.31	18	4	102.60
1	8	91.30	9	15	91.52	18	5	93.64
1	9	101.29	9	16	91.44	18	6	99.30
1	10	99.49	10	1	105.00	18	7	106.71
1	11	94.98	10	2	101.55	18	8	99.93
1	12	96.47	10	3	90.37	18	9	78.13
1	13	102.67	10	4	97.19	18	10	104.01

1	14	108.44	10	5	113.93	18	11	101.57
1	15	122.05	10	6	105.33	18	12	92.52
1	16	103.75	10	7	98.69	18	13	106.56
2	1	93.85	10	8	100.66	18	14	95.86
2	2	100.56	10	9	115.90	18	15	114.54
2	3	96.25	10	10	91.66	18	16	81.00
2	4	105.65	10	11	93.05	19	1	90.33
2	5	97.15	10	12	87.78	19	2	93.36
2	6	88.83	10	13	98.98	19	3	105.93
2	7	103.55	10	14	95.93	19	4	96.63
2	8	100.34	10	15	105.94	19	5	76.83
2	9	93.29	10	16	109.21	19	6	92.26
2	10	105.03	11	1	85.50	19	7	92.96
2	11	95.88	11	2	79.54	19	8	115.51
2	12	103.41	11	3	104.30	19	9	97.44
2	13	104.44	11	4	86.58	19	10	113.60
2	14	118.71	11	5	109.27	19	11	102.93
3	1	112.94	11	6	105.63	19	12	90.69
3	2	111.13	11	7	97.31	19	13	106.52
3	3	100.16	11	8	108.39	19	14	108.91
3	4	104.02	11	9	88.76	19	15	107.53
3	5	88.17	11	10	88.36	20	1	136.43
3	6	88.33	11	11	99.08	20	2	129.34
3	7	105.26	11	12	96.16	20	3	126.68
3	8	82.87	11	13	94.46	20	4	128.53
3	9	105.01	11	14	106.26	20	5	122.20
3	10	97.98	11	15	76.47	20	6	104.48
3	11	97.03	12	1	92.80	20	7	114.37
3	12	94.19	12	2	102.99	20	8	115.21
3	13	110.47	12	3	85.80	20	9	112.32

3	14	103.38	12	4	99.58	20	10	124.51
3	15	99.35	12	5	89.22	20	11	122.92
4	1	90.37	12	6	97.40	20	12	114.55
4	2	110.11	12	7	109.10	20	13	118.72
4	3	125.58	12	8	115.58	20	14	109.61
4	4	123.10	12	9	102.14	20	15	97.85
4	5	94.95	12	10	95.38	21	1	120.40
4	6	99.26	12	11	99.77	21	2	114.88
4	7	114.73	12	12	77.46	21	3	105.97
4	8	115.51	12	13	84.05	21	4	91.91
4	9	110.13	12	14	103.67	21	5	80.86
4	10	83.65	12	15	106.33	21	6	110.17
4	11	85.76	13	1	106.74	21	7	92.99
4	12	94.68	13	2	103.19	21	8	118.10
4	13	100.15	13	3	112.91	21	9	122.37
4	14	101.74	13	4	95.74	21	10	107.15
4	15	97.50	13	5	95.09	21	11	101.76
4	16	106.03	13	6	107.41	21	12	108.80
5	1	107.40	13	7	100.56	21	13	88.80
5	2	86.25	13	8	112.00	21	14	99.81
5	3	98.03	13	9	103.06	21	15	123.54
5	4	106.73	13	10	109.41	21	16	91.59
5	5	115.95	13	11	93.88	22	1	100.55
5	6	114.98	13	12	97.14	22	2	84.49
5	7	85.55	13	13	97.52	22	3	92.34
5	8	109.19	13	14	107.23	22	4	79.45
5	9	109.21	13	15	102.04	22	5	108.18
5	10	102.08	14	1	104.19	22	6	96.46
5	11	94.09	14	2	92.46	22	7	91.34
5	12	95.22	14	3	80.45	22	8	104.91

5	13	96.13	14	4	86.27	22	9	94.95
5	14	105.37	14	5	95.68	22	10	88.45
5	15	89.77	14	6	111.68	22	11	109.73
6	1	97.98	14	7	94.13	22	12	87.40
6	2	93.66	14	8	112.08	22	13	101.71
6	3	95.24	14	9	97.35	22	14	90.14
6	4	72.89	14	10	88.68	22	15	96.65
6	5	95.67	14	11	92.67	22	16	132.15
6	6	110.10	14	12	109.16	23	1	101.96
6	7	106.53	14	13	88.23	23	2	92.75
6	8	100.30	14	14	123.27	23	3	101.63
6	9	88.28	14	15	95.21	23	4	102.23
6	10	105.03	15	1	107.27	23	5	77.82
6	11	99.73	15	2	101.26	23	6	120.48
6	12	105.20	15	3	95.36	23	7	97.02
6	13	81.55	15	4	105.96	23	8	104.01
6	14	93.11	15	5	107.23	23	9	104.21
7	1	91.08	15	6	97.96	23	10	87.75
7	2	97.71	15	7	96.10	23	11	94.02
7	3	96.06	15	8	110.51	23	12	116.87
7	4	98.95	15	9	94.70	23	13	87.49
7	5	112.87	15	10	97.31	23	14	104.47
7	6	96.76	15	11	114.90	23	15	106.57
7	7	109.37	15	12	110.62	23	16	114.11
7	8	103.42	15	13	95.37	24	1	86.54
7	9	95.92	15	14	100.78	24	2	114.14
7	10	85.22	16	1	100.17	24	3	84.84
7	11	104.70	16	2	115.42	24	4	91.94
7	12	114.04	16	3	95.19	24	5	105.17
7	13	92.53	16	4	114.66	24	6	118.29

7	14	104.76	16	5	100.69	24	7	103.80
7	15	107.20	16	6	82.90	24	8	73.50
8	1	102.52	16	7	99.32	24	9	100.23
8	2	105.56	16	8	103.61	24	10	102.05
8	3	118.73	16	9	105.94	24	11	91.96
8	4	89.68	16	10	93.68	24	12	103.00
8	5	96.07	16	11	97.81	24	13	86.11
8	6	109.30	16	12	95.62	24	14	111.74
8	7	88.51	16	13	91.53	24	15	105.26
8	8	93.03	16	14	82.50	25	1	109.27
8	9	74.53	16	15	115.60	25	2	116.91
8	10	109.52	17	1	101.40	25	3	96.12
8	11	95.37	17	2	115.30	25	4	100.35
8	12	99.73	17	3	84.33	25	5	115.03
8	13	119.02	17	4	92.50	25	6	90.59
8	14	104.58	17	5	93.85	25	7	109.71
8	15	91.19	17	6	104.58	25	8	105.22
9	1	106.61	17	7	108.25	25	9	83.47
9	2	85.41	17	8	89.83	25	10	100.21
9	3	91.39	17	9	81.30	25	11	105.34
9	4	113.39	17	10	113.77	25	12	105.64
9	5	99.42	17	11	93.36	25	13	102.10
9	6	110.22	17	12	91.55	25	14	126.13
9	7	102.99	17	13	92.21			

Tabela 2.6: Valores gerados para o conjunto de dados do Exemplo 3.

Amostra	Obs	Capac.	Amostra	Obs	Capac.	Amostra	Obs	Capac.
1	1	110.61	9	2	100.38	17	3	108.26
1	2	95.00	9	3	96.82	17	4	87.45

1	3	108.43	9	4	94.23	18	1	110.46
1	4	95.30	10	1	100.82	18	2	85.89
2	1	104.98	10	2	107.16	18	3	105.59
2	2	99.57	10	3	117.64	18	4	96.91
2	3	107.89	10	4	95.38	19	1	98.20
2	4	101.77	11	1	115.12	19	2	92.18
3	1	98.19	11	2	109.71	19	3	104.78
3	2	112.81	11	3	95.43	19	4	89.48
3	3	89.93	11	4	92.51	20	1	128.01
3	4	90.88	12	1	88.84	20	2	125.08
4	1	107.61	12	2	101.60	20	3	115.05
4	2	100.00	12	3	110.94	20	4	121.90
4	3	106.75	12	4	95.63	21	1	99.32
4	4	102.60	13	1	99.21	21	2	115.93
5	1	87.67	13	2	81.24	21	3	105.33
5	2	103.97	13	3	95.71	21	4	86.10
5	3	93.76	13	4	99.46	22	1	96.94
5	4	87.19	14	1	98.56	22	2	101.09
6	1	92.95	14	2	96.94	22	3	101.23
6	2	96.72	14	3	85.24	22	4	89.22
6	3	90.75	14	4	99.12	23	1	108.28
6	4	90.69	15	1	106.96	23	2	108.89
7	1	100.09	15	2	90.33	23	3	93.15
7	2	98.58	15	3	112.15	23	4	112.45
7	3	97.62	15	4	72.34	24	1	103.45
7	4	83.17	16	1	88.42	24	2	111.58
8	1	101.73	16	2	108.42	24	3	89.72
8	2	112.99	16	3	99.51	24	4	103.02
8	3	105.05	16	4	93.28	25	1	119.01
8	4	90.38	17	1	98.97	25	2	84.17

9	1	97.13	17	2	102.16	25	3	107.81
25	4	85.18						

Tabela 2.7: Valores gerados para o conjunto de dados do Exemplo 4.

Amostra	Capacidade	Amostra	Capacidade
1	99.44	14	94.19
2	106.63	15	120.47
3	85.78	16	103.46
4	101.67	17	83.29
5	110.93	18	109.46
6	122.88	19	100.97
7	116.47	20	108.27
8	87.20	21	101.14
9	101.49	22	119.86
10	109.06	23	109.44
11	105.51	24	115.50
12	108.19	25	119.35
13	113.84		

Tabela 3.3: Valores gerados para o conjunto de dados do Exemplo 5.

Amostra (com 100 aparelhos cada)	Total de aparelhos defeituosos	Amostra (com 100 aparelhos cada)	Total de aparelhos defeituosos
1	14	14	7
2	7	15	9
3	11	16	8
4	13	17	11
5	6	18	9
6	5	19	4
7	10	20	12

8	13	21	18
9	9	22	16
10	9	23	16
11	11	24	18
12	12	25	21
13	8		

Tabela 3.4: Valores gerados para o conjunto de dados do Exemplo 6.

Amostra	Total de aparelhos observados	Total de aparelhos defeituosos
1	101	12
2	94	11
3	108	10
4	127	7
5	113	14
6	137	11
7	136	14
8	90	6
9	137	17
10	93	10
11	130	15
12	130	14
13	125	7
14	129	12
15	140	17
16	96	13
17	98	6
18	140	14
19	104	15
20	137	16

21	139	22
22	111	18
23	111	21
24	137	18
25	119	20

Tabela 4.3: Valores gerados para o conjunto de dados do Exemplo 8.

Amostra (com 100 aparelhos cada)	Total de defeitos na amostra	Amostra (com 100 aparelhos cada)	Total de defeitos na amostra
1	13	14	15
2	10	15	9
3	10	16	12
4	12	17	9
5	5	18	9
6	7	19	8
7	2	20	13
8	6	21	20
9	12	22	15
10	6	23	21
11	14	24	22
12	9	25	17
13	9		

Tabela 4.4: Valores gerados para o conjunto de dados do Exemplo 9.

Amostra	Sub-amostra	Número de defeitos	Amostra	Sub-amostra	Número de defeitos	Amostra	Sub-amostra	Número de defeitos
1	1	14	9	3	8	17	5	9
1	2	8	9	4	13	18	1	12
1	3	12	9	5	6	18	2	12

1	4	16	10	1	13	18	3	15
1	5	15	10	2	8	18	4	9
2	1	6	10	3	9	18	5	10
2	2	9	10	4	12	19	1	9
2	3	11	10	5	10	19	2	8
2	4	10	11	1	10	19	3	12
2	5	7	11	2	13	19	4	7
3	1	6	11	3	5	19	5	8
3	2	7	11	4	12	20	1	15
3	3	14	11	5	12	20	2	10
3	4	5	12	1	6	20	3	10
3	5	6	12	2	5	20	4	10
4	1	14	12	3	7	20	5	11
4	2	13	12	4	5	21	1	9
4	3	5	12	5	13	21	2	13
4	4	13	13	1	9	21	3	17
4	5	9	13	2	12	21	4	20
5	1	12	13	3	9	21	5	11
5	2	13	13	4	14	22	1	9
5	3	13	13	5	8	22	2	18
5	4	14	14	1	9	22	3	18
5	5	9	14	2	11	22	4	24
6	1	16	14	3	9	22	5	13
6	2	10	14	4	17	23	1	14
6	3	10	14	5	9	23	2	14
6	4	19	15	1	9	23	3	18
6	5	7	15	2	12	23	4	13
7	1	10	15	3	9	23	5	14
7	2	13	15	4	8	24	1	11
7	3	9	15	5	7	24	2	15

7	4	10	16	1	8	24	3	20
7	5	10	16	2	8	24	4	14
8	1	8	16	3	14	24	5	11
8	2	12	16	4	6	25	1	11
8	3	6	16	5	7	25	2	11
8	4	7	17	1	10	25	3	13
8	5	11	17	2	9	25	4	11
9	1	6	17	3	6	25	5	20
9	2	10	17	4	9			

Tabela 4.5: Valores gerados para o conjunto de dados do Exemplo 10.

Amostra	Sub-amostra	Número de defeitos	Amostra	Sub-amostra	Número de defeitos	Amostra	Sub-amostra	Número de defeitos
1	1	12	10	1	5	17	6	11
1	2	8	10	2	8	18	1	14
1	3	7	10	3	7	18	2	10
1	4	9	10	4	12	18	3	8
2	1	12	10	5	8	18	4	10
2	2	9	10	6	13	18	5	10
2	3	4	10	7	11	18	6	16
2	4	5	11	1	13	19	1	8
3	1	5	11	2	11	19	2	15
3	2	12	11	3	17	19	3	13
3	3	6	11	4	9	19	4	11
3	4	17	11	5	9	19	5	7
3	5	12	12	1	12	20	1	7
3	6	10	12	2	12	20	2	13
3	7	11	12	3	4	20	3	4
4	1	10	12	4	12	20	4	4

4	2	5	13	1	13	20	5	14
4	3	12	13	2	7	20	6	16
4	4	8	13	3	9	21	1	14
4	5	8	13	4	9	21	2	12
4	6	8	13	5	17	21	3	15
5	1	11	13	6	11	21	4	12
5	2	11	14	1	13	22	1	15
5	3	10	14	2	9	22	2	19
5	4	12	14	3	16	22	3	11
5	5	5	14	4	9	22	4	14
5	6	7	15	1	16	22	5	10
6	1	14	15	2	11	22	6	16
6	2	7	15	3	13	23	1	20
6	3	14	15	4	5	23	2	19
6	4	12	15	5	10	23	3	14
7	1	14	16	1	7	23	4	14
7	2	10	16	2	8	23	5	10
7	3	9	16	3	6	23	6	12
7	4	5	16	4	4	23	7	14
8	1	5	16	5	11	24	1	17
8	2	11	16	6	7	24	2	19
8	3	18	16	7	4	24	3	13
8	4	7	17	1	8	24	4	14
9	1	10	17	2	15	25	1	15
9	2	13	17	3	14	25	2	11
9	3	12	17	4	8	25	3	25
9	4	15	17	5	7	25	4	22

Tabela 5.3: Valores gerados para o conjunto de dados do Exemplo 11.

Amostra	Obs	Capac.	Amostra	Obs	Capac.	Amostra	Obs	Capac.
1	1	101.2	9	6	89.98	17	11	87.83
1	2	81.88	9	7	105.49	17	12	88.22
1	3	101.52	9	8	98.6	17	13	83.73
1	4	88.81	9	9	103.76	17	14	103.11
1	5	100.02	9	10	116.48	17	15	97.63
1	6	111.89	9	11	101.01	18	1	106.59
1	7	94.95	9	12	106.11	18	2	87.04
1	8	99.01	9	13	109.68	18	3	113.6
1	9	103.05	9	14	102.82	18	4	121.44
1	10	101.98	9	15	102.29	18	5	94.28
1	11	99.51	10	1	90.53	18	6	92.31
1	12	91.57	10	2	98.58	18	7	121.6
1	13	79.25	10	3	102.13	18	8	108.37
1	14	96.39	10	4	103.12	18	9	106.22
1	15	93.62	10	5	78.51	18	10	94.09
2	1	96.34	10	6	85.55	18	11	84.81
2	2	123.55	10	7	103.87	18	12	94.62
2	3	110.93	10	8	93.22	18	13	112.06
2	4	102.86	10	9	76.43	18	14	102.23
2	5	109.94	10	10	88.97	18	15	89.06
2	6	84.81	10	11	105.48	19	1	105.78
2	7	114.97	10	12	123.13	19	2	110.41
2	8	108.2	10	13	87.49	19	3	92.39
2	9	110.66	10	14	93.62	19	4	105.49
2	10	107.34	10	15	102.23	19	5	108.28
2	11	109.6	11	1	83.3	19	6	112.26
2	12	93.08	11	2	90.41	19	7	92.94
2	13	114.06	11	3	111.18	19	8	97.78

2	14	83.66	11	4	101.44	19	9	101.25
2	15	102.62	11	5	102.19	19	10	110.31
3	1	115.65	11	6	90.14	19	11	106.24
3	2	103.15	11	7	123.61	19	12	98.9
3	3	90.65	11	8	92.63	19	13	90.43
3	4	98.75	11	9	92.4	19	14	103.85
3	5	94.73	11	10	128.61	19	15	96
3	6	98.33	11	11	100.57	20	1	101.74
3	7	96.91	11	12	113.06	20	2	89.6
3	8	99.46	11	13	99.62	20	3	109.41
3	9	104.91	11	14	91.67	20	4	113.22
3	10	106.39	11	15	95.3	20	5	105.33
3	11	111.51	12	1	96.99	20	6	103.72
3	12	97.13	12	2	98.43	20	7	107.31
3	13	92.55	12	3	95.28	20	8	94.49
3	14	109.2	12	4	105.22	20	9	99.53
3	15	107.81	12	5	96.42	20	10	99.75
4	1	106.27	12	6	96.36	20	11	107.69
4	2	104.27	12	7	95.52	20	12	92.74
4	3	86.87	12	8	94.33	20	13	89.66
4	4	98.32	12	9	85.36	20	14	91.43
4	5	84.82	12	10	98.15	20	15	101.28
4	6	88.37	12	11	98.39	21	1	77.9
4	7	103.27	12	12	110.63	21	2	95.41
4	8	89.71	12	13	89.18	21	3	94.17
4	9	93.17	12	14	91.14	21	4	94.42
4	10	98.4	12	15	105.08	21	5	97.46
4	11	102.23	13	1	81.89	21	6	113.77
4	12	105.28	13	2	116.42	21	7	99.68
4	13	85.89	13	3	96.58	21	8	111.05

4	14	83.67	13	4	84.05	21	9	89.81
4	15	95.12	13	5	91.05	21	10	100
5	1	115.65	13	6	112.18	21	11	100.35
5	2	100.53	13	7	107.43	21	12	109.56
5	3	114.38	13	8	102.08	21	13	102.45
5	4	105.06	13	9	98.27	21	14	115.81
5	5	92.66	13	10	115.55	21	15	87.66
5	6	99.54	13	11	105.76	22	1	96.03
5	7	105.52	13	12	78.66	22	2	99.42
5	8	101.38	13	13	90.81	22	3	107.63
5	9	84.81	13	14	95.57	22	4	99.34
5	10	103.82	13	15	104.37	22	5	91.31
5	11	97.16	14	1	101.68	22	6	106.17
5	12	84.44	14	2	107.64	22	7	101.95
5	13	92.26	14	3	117.27	22	8	112.1
5	14	100.35	14	4	114.22	22	9	109.65
5	15	102.11	14	5	99.06	22	10	102.77
6	1	104.12	14	6	104.94	22	11	113.42
6	2	92.78	14	7	116.07	22	12	126.01
6	3	109.04	14	8	108.69	22	13	103.26
6	4	107.04	14	9	95.71	22	14	91.25
6	5	94.18	14	10	111.86	22	15	96.17
6	6	96.06	14	11	108.24	23	1	98.66
6	7	99.69	14	12	107.16	23	2	99.91
6	8	86.94	14	13	91.64	23	3	104.45
6	9	81.76	14	14	104.6	23	4	99.22
6	10	123.68	14	15	103.41	23	5	113.22
6	11	102.15	15	1	88.53	23	6	104.21
6	12	98.11	15	2	117.59	23	7	96.27
6	13	118.07	15	3	94.94	23	8	103.1

6	14	108.83	15	4	111.44	23	9	107.89
6	15	103.84	15	5	105.03	23	10	88.58
7	1	94.44	15	6	97.08	23	11	93.86
7	2	103.6	15	7	69.28	23	12	94.78
7	3	99.15	15	8	111.3	23	13	108.64
7	4	86.36	15	9	93.46	23	14	95.69
7	5	96.42	15	10	87.06	23	15	94.05
7	6	92.45	15	11	87.25	24	1	74.94
7	7	80.72	15	12	97.11	24	2	115.13
7	8	107.4	15	13	105.39	24	3	92.01
7	9	102.33	15	14	88.97	24	4	90.9
7	10	109.91	15	15	113.31	24	5	88.04
7	11	85.28	16	1	88.18	24	6	91.82
7	12	93.5	16	2	93.27	24	7	103.8
7	13	103.62	16	3	91.45	24	8	117.97
7	14	112.64	16	4	87.96	24	9	101.84
7	15	104.49	16	5	106.44	24	10	83.52
8	1	115.06	16	6	93.44	24	11	115.1
8	2	100.57	16	7	104.89	24	12	101.96
8	3	109.21	16	8	103.34	24	13	105.07
8	4	100.5	16	9	89.03	24	14	106.83
8	5	115.97	16	10	99.96	24	15	101.68
8	6	101.83	16	11	99.18	25	1	114.05
8	7	113.44	16	12	88.56	25	2	98.38
8	8	117.54	16	13	88.28	25	3	88.64
8	9	90.01	16	14	80.67	25	4	94.79
8	10	84.02	16	15	101.73	25	5	93.08
8	11	109.73	17	1	106.73	25	6	89.43
8	12	101.29	17	2	100.28	25	7	95.44
8	13	104.23	17	3	87.44	25	8	91.1

8	14	105.53	17	4	99.93	25	9	100.43
8	15	92.7	17	5	93.86	25	10	103.85
9	1	98.53	17	6	108.08	25	11	102.62
9	2	95.78	17	7	107.15	25	12	89.79
9	3	105.4	17	8	118.06	25	13	97.1
9	4	103.63	17	9	89.35	25	14	96.29
9	5	81.75	17	10	115.81	25	15	101.8

Tabela 6.3: Valores gerados para o conjunto de dados do Exemplo 12.

Amostra	Obs	V1	V2	V3
1	1	100.84	99.29	99.19
1	2	99.15	98.96	100.35
1	3	98.25	99.63	100.17
1	4	98.31	100.83	97.84
1	5	99.99	100.50	101.34
1	6	100.04	99.77	101.77
1	7	97.94	99.50	98.64
1	8	101.96	100.41	99.60
1	9	99.85	99.95	99.77
1	10	99.89	100.43	100.36
1	11	101.31	101.11	102.08
1	12	100.23	99.07	101.93
2	1	99.64	100.52	99.60
2	2	100.11	100.49	101.51
2	3	99.88	100.56	102.63
2	4	99.20	98.89	97.85
2	5	100.35	100.46	100.33
2	6	99.85	98.63	100.48
2	7	100.30	100.15	100.57

2	8	100.78	98.82	100.25
2	9	100.02	100.13	101.85
2	10	101.74	101.47	101.10
2	11	98.90	98.59	97.66
2	12	100.06	99.81	100.96
3	1	100.08	100.81	100.13
3	2	99.00	98.62	99.07
3	3	99.75	99.89	99.33
3	4	99.48	100.20	100.88
3	5	98.66	100.49	99.75
3	6	101.47	100.75	99.30
3	7	100.32	97.93	99.29
3	8	101.32	101.90	99.37
3	9	99.38	100.79	100.21
3	10	100.15	98.17	101.02
3	11	100.11	100.22	99.95
3	12	99.25	100.21	100.41
4	1	98.07	98.15	99.28
4	2	99.72	101.08	99.27
4	3	100.27	100.09	99.52
4	4	99.75	99.26	98.61
4	5	102.39	101.64	99.76
4	6	100.34	100.40	100.75
4	7	100.48	101.08	100.61
4	8	99.72	99.73	99.55
4	9	100.36	99.54	101.02
4	10	99.52	99.77	99.19
4	11	99.06	100.03	100.59
4	12	98.37	100.19	99.56
5	1	100.25	99.51	101.00

5	2	101.15	102.15	100.00
5	3	99.99	99.27	101.02
5	4	99.51	98.29	99.92
5	5	99.74	100.44	97.21
5	6	99.06	100.21	99.15
5	7	100.23	100.68	99.50
5	8	99.44	98.95	100.79
5	9	99.97	100.85	100.35
5	10	100.60	99.13	101.31
5	11	100.17	101.46	98.76
5	12	100.99	100.30	100.73
6	1	101.73	99.57	100.71
6	2	99.69	100.82	100.70
6	3	99.28	100.33	100.41
6	4	100.47	99.78	100.64
6	5	99.68	99.01	100.64
6	6	101.65	100.58	100.98
6	7	101.12	101.01	99.13
6	8	98.98	99.52	98.55
6	9	98.31	99.56	99.93
6	10	100.09	100.47	100.00
6	11	99.96	100.12	99.30
6	12	99.81	100.01	99.60
7	1	100.25	100.80	100.90
7	2	100.22	99.03	100.84
7	3	99.45	99.66	100.42
7	4	98.37	100.71	97.93
7	5	100.74	98.92	98.97
7	6	97.99	99.66	98.87
7	7	100.64	101.95	101.47

7	8	99.83	100.03	101.75
7	9	100.93	98.33	99.78
7	10	101.65	101.80	100.93
7	11	99.98	100.94	100.36
7	12	100.35	99.87	100.57
8	1	99.55	98.33	100.96
8	2	99.77	99.45	100.27
8	3	100.62	99.43	99.73
8	4	98.39	98.14	97.23
8	5	100.32	99.10	98.55
8	6	100.13	99.83	99.79
8	7	99.92	100.87	99.89
8	8	98.47	98.92	97.51
8	9	100.94	101.03	101.00
8	10	98.98	99.14	99.65
8	11	98.87	99.96	99.16
8	12	99.44	100.20	99.26
9	1	100.27	101.59	101.87
9	2	99.52	98.53	101.82
9	3	99.37	100.29	99.37
9	4	101.14	99.95	99.64
9	5	100.29	100.26	101.92
9	6	100.35	100.92	100.87
9	7	99.52	99.84	98.93
9	8	100.16	100.53	100.82
9	9	100.47	100.11	100.22
9	10	99.36	99.69	99.96
9	11	98.10	100.73	99.67
9	12	99.84	100.00	99.42
10	1	99.43	99.88	98.49

10	2	99.22	98.32	101.04
10	3	97.54	97.73	98.58
10	4	99.54	99.70	99.16
10	5	100.62	100.95	99.45
10	6	99.92	100.70	99.38
10	7	100.45	100.93	99.52
10	8	99.67	98.63	98.15
10	9	99.56	99.90	101.32
10	10	100.96	99.95	99.86
10	11	100.19	98.74	100.05
10	12	100.66	100.94	100.96
11	1	99.02	100.44	100.47
11	2	101.05	100.72	101.05
11	3	99.56	99.00	100.05
11	4	101.39	99.73	100.29
11	5	100.37	98.47	100.02
11	6	99.09	99.76	99.90
11	7	99.98	100.20	101.50
11	8	98.16	99.84	99.26
11	9	100.29	100.49	98.81
11	10	99.44	99.48	97.72
11	11	101.21	100.41	98.99
11	12	98.72	100.17	99.19
12	1	99.19	98.85	98.71
12	2	100.00	99.18	99.01
12	3	98.21	99.01	98.64
12	4	99.01	100.44	100.98
12	5	100.37	100.41	100.24
12	6	99.87	101.00	101.58
12	7	100.63	101.01	101.63

12	8	101.66	101.59	101.15
12	9	100.70	99.47	101.31
12	10	100.48	99.54	100.56
12	11	99.95	101.40	98.17
12	12	99.44	98.71	99.41
13	1	99.63	99.57	99.19
13	2	100.46	101.19	101.20
13	3	100.09	101.16	101.23
13	4	100.92	99.81	100.45
13	5	100.83	100.49	101.36
13	6	98.85	100.45	101.02
13	7	99.67	100.62	99.76
13	8	99.23	98.16	96.97
13	9	99.77	99.90	99.68
13	10	99.65	99.13	99.70
13	11	100.67	101.71	100.93
13	12	99.93	100.96	99.66
14	1	99.37	99.61	99.18
14	2	99.29	99.38	99.42
14	3	99.76	100.28	102.17
14	4	100.02	99.50	99.22
14	5	99.37	100.65	98.65
14	6	100.92	100.78	102.08
14	7	99.86	101.06	99.49
14	8	99.51	101.06	100.89
14	9	100.66	100.20	100.89
14	10	100.31	99.95	101.67
14	11	98.75	99.82	100.77
14	12	101.11	101.19	101.20
15	1	100.16	100.46	99.47

15	2	101.33	101.20	100.26
15	3	99.53	97.21	98.58
15	4	98.23	100.35	100.78
15	5	98.53	98.32	98.67
15	6	99.98	100.99	100.25
15	7	99.92	99.53	99.55
15	8	100.38	100.04	99.03
15	9	100.84	102.28	101.29
15	10	99.74	100.42	100.71
15	11	99.56	99.27	101.05
15	12	100.07	100.64	100.11
16	1	100.88	99.73	101.56
16	2	99.88	99.63	100.02
16	3	99.87	99.81	100.06
16	4	99.57	99.83	98.77
16	5	99.69	100.63	100.59
16	6	99.57	99.43	99.28
16	7	99.45	99.41	99.84
16	8	99.41	97.83	98.44
16	9	100.98	98.89	99.83
16	10	99.84	99.62	99.99
16	11	100.65	99.18	99.92
16	12	100.02	101.10	99.77
17	1	99.40	99.91	100.62
17	2	100.00	98.52	98.43
17	3	100.64	99.87	100.30
17	4	100.85	100.49	99.78
17	5	100.44	99.38	100.11
17	6	99.72	99.24	99.38
17	7	99.77	100.21	101.57

17	8	101.10	100.99	99.91
17	9	98.98	99.26	98.51
17	10	100.77	99.28	100.70
17	11	100.08	98.74	100.02
17	12	99.59	99.93	101.11
18	1	100.59	99.14	100.95
18	2	97.82	98.78	98.86
18	3	100.50	100.73	100.11
18	4	98.66	99.13	98.86
18	5	98.89	99.46	99.19
18	6	98.33	98.93	98.33
18	7	99.84	100.67	100.27
18	8	100.49	99.72	99.75
18	9	101.89	100.85	99.74
18	10	100.36	100.16	99.95
18	11	101.17	101.82	100.27
18	12	99.30	100.28	99.58
19	1	101.52	100.80	101.03
19	2	101.90	102.81	102.43
19	3	100.07	100.20	100.70
19	4	101.61	101.74	101.57
19	5	99.91	99.10	98.35
19	6	101.19	101.04	101.69
19	7	99.48	101.57	99.75
19	8	99.47	100.35	99.77
19	9	99.90	99.05	100.08
19	10	99.12	98.70	99.36
19	11	100.96	100.82	100.68
19	12	100.63	100.75	101.25
20	1	98.34	100.92	98.30

20	2	98.91	99.92	99.92
20	3	100.08	99.99	98.93
20	4	99.28	98.56	99.42
20	5	100.97	101.16	100.15
20	6	98.74	99.57	99.16
20	7	99.30	98.59	99.50
20	8	101.32	100.70	101.16
20	9	98.68	99.53	99.75
20	10	100.71	100.00	101.31
20	11	100.23	98.88	100.43
20	12	101.31	101.60	100.97
21	1	100.00	101.65	101.48
21	2	98.72	100.11	100.98
21	3	100.70	99.91	102.35
21	4	100.53	101.49	102.97
21	5	99.37	99.93	101.43
21	6	100.28	100.08	101.35
21	7	99.46	99.89	101.55
21	8	100.10	99.71	101.88
21	9	98.77	99.76	101.67
21	10	100.63	99.74	102.31
21	11	100.38	100.29	102.38
21	12	100.31	100.42	101.01
22	1	101.84	99.23	103.52
22	2	100.37	99.63	100.59
22	3	100.78	100.65	102.41
22	4	100.10	100.24	102.33
22	5	100.12	99.94	103.09
22	6	99.54	98.38	101.21
22	7	100.47	98.88	102.66

22	8	99.83	98.19	101.26
22	9	99.70	98.88	102.55
22	10	100.40	99.10	102.36
22	11	101.73	100.83	103.51
22	12	100.87	100.64	103.90
23	1	99.32	100.45	100.25
23	2	100.01	100.41	101.32
23	3	99.05	98.95	101.56
23	4	99.37	100.10	101.68
23	5	98.91	99.74	101.96
23	6	99.01	99.12	100.63
23	7	101.13	99.85	102.81
23	8	101.01	101.12	104.51
23	9	99.05	100.04	102.86
23	10	100.10	101.88	103.36
23	11	100.87	98.87	100.97
23	12	99.41	101.52	103.00
24	1	100.99	100.97	101.95
24	2	100.43	99.82	101.32
24	3	98.64	99.19	102.37
24	4	99.30	98.18	101.16
24	5	101.65	99.92	102.02
24	6	101.71	101.45	103.22
24	7	98.65	100.11	100.96
24	8	100.25	99.97	102.46
24	9	99.11	98.32	100.95
24	10	100.36	101.01	101.41
24	11	99.21	99.91	100.09
24	12	100.77	100.03	102.17
25	1	101.59	100.86	102.32

25	2	100.46	100.36	102.17
25	3	100.53	99.50	101.50
25	4	101.17	100.16	103.09
25	5	101.98	101.91	102.63
25	6	100.60	100.69	102.43
25	7	99.62	101.45	102.27
25	8	101.03	100.33	103.09
25	9	100.45	101.49	103.33
25	10	101.55	100.58	102.95
25	11	98.56	98.29	101.50
25	12	101.97	101.83	103.69

Tabela 6.4: Valores gerados para o conjunto de dados do Exemplo 13.

Amostra	V1	V2	V3
1	98.37	99.56	97.46
2	100.56	100.07	100.53
3	101.24	99.76	100.73
4	99.21	100.95	100.62
5	99.05	99.84	99.42
6	99.78	99.52	99.37
7	98.96	99.64	100.00
8	99.65	99.90	99.01
9	98.05	99.89	99.97
10	98.68	101.08	98.70
11	100.53	100.33	100.64
12	99.08	99.42	98.30
13	99.71	100.28	99.87
14	101.07	101.43	100.18
15	100.49	101.20	101.65

16	99.26	98.24	99.29
17	100.67	99.29	99.84
18	99.95	101.74	100.70
19	100.10	99.64	100.75
20	99.37	100.93	100.59
21	99.10	99.32	102.53
22	101.81	99.63	103.44
23	99.86	100.94	103.43
24	100.49	99.66	102.96
25	99.68	98.22	102.46

Apêndice C

Neste apêndice estão expostos os algoritmos para construção e análise da maioria dos gráficos de controle discutidos.

Quadro 2.1.a: Algoritmo para controle da variabilidade e da média, usando os Gráficos S e \bar{X} .

A – Procedimentos iniciais

- 1- Estabelecer o número de amostras m . Sugestão: use m em torno de 25.
- 2- Definir o tamanho de cada amostra n_i , considerando que o Gráfico S é mais recomendado nos casos em que $n_i > 10$ ou 12.
- 3- Considerar i para a amostra e j para as observações. Coletar então os valores da variável de interesse (x_{ij}) na indústria ou estabelecimento em questão.
- 4- Calcular o desvio-padrão amostral para cada i , através de

$$S_i = \sqrt{\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2}.$$

- 5- Calcular a média amostral para cada i , através de $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$.

B – Construindo os gráficos de acordo com suas características

B.1 – As m amostras são de mesmo tamanho

- 1- Construir o gráfico de S_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores de desvio-padrão calculados de cada amostra.
- 2- Calcular o valor de $c_4 = \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]} \sqrt{\frac{2}{n-1}}$.
- 3- Grafar o valor de referência, que é dado por $c_4\sigma$ se houver especificação para o desvio-padrão do processo, ou $\bar{S} = \frac{1}{m} \sum_{i=1}^m S_i$, que é a média dos desvios-padrão, caso não haja especificação e a variabilidade precise ser estimada.
- 4- Calcular e grafar os limites de controle, dados por $c_4\sigma \pm 3\sigma\sqrt{(1-c_4^2)}$ caso haja

especificação, ou $\bar{S} \pm 3 \frac{\bar{S}}{c_4} \sqrt{(1-c_4^2)}$ caso seja necessário estimar o parâmetro.

5- Construir o gráfico de \bar{x}_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da média de cada amostra.

6- Grafar o valor de referência, que é dado por μ^0 se houver especificação ou $\bar{\bar{x}} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i$, que é a média das médias, caso não haja especificação e a média do processo precise ser estimada.

7- Grafar os limites de controle $\mu^0 \pm 3 \frac{\sigma}{\sqrt{n}}$, caso haja especificações ou

$\bar{\bar{x}} \pm 3 \frac{\bar{S}}{c_4 \sqrt{n}}$, caso os parâmetros precisem ser estimados.

B.2 – As m amostras são de tamanhos variáveis

1- Construir o gráfico de S_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores de desvio-padrão calculados de cada amostra.

2- Calcular, para cada amostra, o valor de $c_{4i} = \frac{\Gamma(n_i/2)}{\Gamma[(n_i-1)/2]} \sqrt{\frac{2}{n_i-1}}$.

3- Grafar o valor de referência, que é dado por $c_{4i} \sigma$ se houver especificação ou

$\bar{S} = \left[\frac{\sum_{i=1}^m (n_i - 1) S_i^2}{\left(\sum_{i=1}^m n_i \right) - m} \right]^{1/2}$, que é a média ponderada dos desvios-padrão para

amostras de tamanhos variáveis, caso não haja especificação e a variabilidade precise ser estimada.

4- Calcular e grafar os limites de controle $c_{4i} \sigma \pm 3 \sigma \sqrt{(1-c_{4i}^2)}$, caso haja especificação, ou $\bar{S} \pm 3 \frac{\bar{S}}{c_{4i}} \sqrt{(1-c_{4i}^2)}$, caso o desvio-padrão do processo precise ser estimado.

5- Construir o gráfico de \bar{x}_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da média de cada amostra.

6- Grafar o valor de referência, que é dado por μ^0 se houver especificação ou

$$\bar{\bar{x}} = \frac{\sum_{i=1}^m n_i \bar{x}_i}{\sum_{i=1}^m n_i}, \text{ que é a média ponderada das médias para amostras de tamanhos}$$

variáveis, caso não haja especificação e a média do processo precise ser estimada.

7- Grafar, para cada amostra, os limites de controle $\mu^0 \pm 3 \frac{\sigma}{\sqrt{n_i}}$, caso haja

$$\text{especificações ou } \bar{\bar{x}} \pm 3 \frac{\bar{S}}{c_{4i} \sqrt{n_i}}, \text{ caso seja necessário estimar os parâmetros da}$$

distribuição da variável observada.

C – Análise dos gráficos

1- Para cada um dos gráficos construídos, verificar a existência dos três possíveis (principais) itens a seguir, que são sinais de falta de controle do processo:

- Existem pontos fora dos limites de controle, isto é, valores da estatística que estão acima do limite superior de controle (*LSC*) ou abaixo do limite inferior de controle (*LIC*)?
- Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo, isto é, há vários pontos consecutivos acima ou abaixo da em um intervalo de tempo?
- Existem tendências ou sazonalidades (padrões de repetição) nos gráficos construídos a partir de amostras de mesmo tamanho?

2- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.

3- É importante ressaltar também que, quando detectada alguma das características acima e uma causa tiver sido identificada como responsável pela falta de controle, o próximo passo é eliminar as amostras que possam ter sofrido efeito dessa causa e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou se existe um problema crônico ou referências (especificidades) fora do ideal.

Quadro 2.2.a: Algoritmo para controle da variabilidade e da média, usando os Gráficos R e \bar{X} .

A – Procedimentos iniciais

- 1- Estabelecer o número de amostras m . Sugestão: use m em torno de 25.
- 2- Definir o tamanho de cada amostra n (n constante), considerando que o Gráfico R é mais apropriado para os casos em que $n < 10$ ou 12.
- 3- Considerar i para a amostra e j para as observações. Coletar então os valores da variável de interesse (x_{ij}) na indústria ou estabelecimento em questão.
- 4- Calcular a amplitude amostral para cada i , através de $R_i = \max_j (x_{ij}) - \min_j (x_{ij})$ caso $2 < n < 10$ ou 12, ou $R_i = |x_i - x_{i-1}|$ se $n=1$.

- 5- Calcular a média amostral para cada i , através de $\bar{x}_i = \frac{\sum_{j=1}^n x_{ij}}{n}$.

B – Construindo os gráficos de acordo com suas características

B.1 – A média e a variabilidade são especificadas

- 1- Construir o gráfico de R_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da amplitude amostral de cada amostra.
- 2- Verificar os valores de d_2 e d_3 , de acordo com o tamanho das amostras.
- 3- Grafar o valor de referência, dado por $d_2\sigma$.
- 4- Calcular e grafar os limites de controle $d_2\sigma \pm 3d_3\sigma$.
- 5- Construir o gráfico de \bar{x}_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da média de cada amostra.
- 6- Grafar o valor de referência, que é dado por μ^0 .
- 7- Grafar os limites de controle $\mu^0 \pm 3\frac{\sigma}{\sqrt{n}}$.

B.2 – A média e a variabilidade não são especificadas

- 1- Construir o gráfico de R_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da amplitude amostral de cada amostra.
- 2- Calcular a amplitude média das amostras, dada por $\bar{R} = \frac{1}{m} \sum_{i=1}^m R_i$ se $n > 1$, ou

$$\bar{R} = \frac{1}{m-1} \sum_{i=2}^m R_i \text{ se } n = 1.$$

- 3- Grafar o valor de referência, dado por \bar{R} .
- 4- Calcular e grafar os limites de controle $\bar{R} \pm 3 \frac{d_3 \bar{R}}{d_2}$.
- 5- Construir o gráfico de \bar{x}_i , tal que na abscissa estejam colocados os indicadores de cada amostra, de 1 a m , e na ordenada os valores da média de cada amostra.
- 6- Grafar o valor de referência, dado por $\bar{\bar{x}} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i$, que é a média das médias.
- 7- Grafar os limites de controle $\bar{\bar{x}} \pm 3 \frac{\bar{R}}{d_2 \sqrt{n}}$.

C – Análise dos gráficos

- 1- Para cada um dos gráficos construídos, verificar a existência dos três possíveis (principais) itens a seguir, que são sinais de falta de controle do processo:
 - Existem pontos fora dos limites de controle, isto é, pontos localizados acima do *LSC* ou abaixo do *LIC*?
 - Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo, isto é, há vários pontos consecutivos acima ou abaixo da linha central do gráfico em um intervalo de tempo?
 - Existem tendências ou sazonalidades?
- 2- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.
- 3- É importante ressaltar novamente que, quando detectada alguma das características acima e uma causa tiver sido identificada como responsável pela falta de controle, o próximo passo é eliminar as amostras que possam ter sofrido efeito dessa causa e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou se existe um problema crônico ou referências (especificidades) fora do ideal.

Quadro 3.1.a: Algoritmo de construção e análise do Gráfico p para tamanhos iguais ou variáveis de amostra.

A – Procedimentos iniciais

- 1- Estabelecer o número de amostras m .
- 2- Definir o tamanho de cada amostra n_i .
- 3- Coletar então as amostras na indústria ou estabelecimento em questão.
- 4- Calcular a proporção de itens defeituosos para cada amostra, isto é, $\hat{p}_i = \frac{D_i}{n_i}$,

em que D_i é o número de itens defeituosos na amostra i .

B – Construindo os gráficos de acordo com suas características

B.1 – As m amostras são de mesmo tamanho

- 1- Construir o gráfico de \hat{p}_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores das proporções de itens defeituosos para cada amostra.
- 2- Grafar o valor de referência, que é dado por p , caso tenha sido especificado um valor ideal p para a proporção de itens defeituosos, ou $\bar{p} = \frac{1}{m} \sum_{i=1}^m \hat{p}_i$, caso contrário.
- 3- Grafar os limites de controle $p \pm 3\sqrt{\frac{p(1-p)}{n}}$, se p tiver sido especificado, ou $\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$, caso contrário.

B.2 – As m amostras são de tamanhos variáveis

- 1- Construir o gráfico de \hat{p}_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores das proporções de itens defeituosos para cada amostra.
- 2- Grafar o valor de referência, que é dado por p , caso tenha sido especificado um valor ideal p para a proporção de itens defeituosos, ou $\bar{p} = \frac{\sum_{i=1}^m D_i}{\sum_{i=1}^m n_i}$, caso contrário.

3- Grafar os limites de controle $p \pm 3\sqrt{\frac{p(1-p)}{n_i}}$, se p tiver sido especificado, ou

$$\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}, \text{ caso contrário.}$$

4- Calcular para cada amostra o valor de z_i , considerando que $z_i = \frac{\hat{p}_i - p}{\sqrt{\frac{p(1-p)}{n_i}}}$,

$$\text{caso } p \text{ tenha sido especificado, ou } z_i = \frac{\hat{p}_i - \bar{p}}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}}, \text{ caso contrário.}$$

5- Construir o gráfico de z_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores de z_i calculados para cada amostra.

6- Grafar o valor de referência, que deve ser igual a 0 e os limites de controle, que devem ser iguais a ± 3 .

C – Análise dos gráficos

1- Verificar a existência dos três possíveis (principais) itens a seguir no(s) gráfico(s) construído(s), que são sinais de falta de controle do processo:

- Existem pontos fora dos limites de controle, isto é, pontos acima do limite superior de controle (*LSC*) ou abaixo do limite inferior de controle (*LIC*)?
- Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo, isto é, há sequências de pontos acima ou abaixo da linha central do gráfico em um intervalo de tempo?
- Existem tendências ou sazonalidades no Gráfico p para amostras de mesmo tamanho ou no Gráfico p padronizado?

2- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.

3- Se uma especificação tiver sido dada e se pelo menos uma das características do item 1 for detectada, é necessário construir o gráfico novamente, mas agora sem as especificações, para verificar se o processo está realmente fora de controle,

ou não consegue suprir as mesmas.

- 4- É importante ressaltar também que, quando detectada alguma característica do item 1 e o item 3 já tiver sido verificado, o próximo passo é eliminar as amostras que apresentaram esse item e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou existe um problema crônico e, assim, poder usar esses limites no futuro.

Quadro 3.2.a: Algoritmo de construção e análise do Gráfico np para tamanhos iguais de amostra.

A – Procedimentos iniciais

- 1- Estabelecer o número de amostras m .
- 2- Definir o tamanho de cada amostra n .
- 3- Coletar então as amostras na indústria ou estabelecimento em questão.
- 4- Calcular o número de itens defeituosos, $n\hat{p}_i = \sum_{j=1}^n x_{ij} = D_i$, para cada amostra.

B – Construindo os gráficos de acordo com suas características

B.1 – Parâmetro especificado (amostras de mesmo tamanho)

- 1- Construir o gráfico de $n\hat{p}_i$, no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores do número de itens defeituosos para cada amostra.
- 2- Grafar a referência, cujo valor é dado por np .
- 3- Grafar os limites de controle $np \pm 3\sqrt{np(1-p)}$.

B.2 – Parâmetro não especificado (amostras de mesmo tamanho)

- 1- Construir o gráfico de $n\hat{p}_i$, no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores do número de itens defeituosos para cada amostra.
- 2- Grafar a referência, cujo valor é dado por $n\bar{p} = \frac{1}{m} \sum_{i=1}^m n\hat{p}_i$.
- 3- Grafar os limites de controle $n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}$.

C – Análise dos gráficos

- 5- Verificar a existência dos três possíveis (principais) itens a seguir no gráfico

construído, que caracterizam a falta de controle do processo:

- Existem pontos fora dos limites de controle, isto é, pontos acima do limite superior de controle (*LSC*) ou abaixo do limite inferior de controle (*LIC*)?
- Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo, isto é, há sequências de pontos acima ou abaixo da linha central do gráfico em um intervalo de tempo?
- Existem tendências ou sazonalidades?

- 1- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.
- 2- Se uma especificação tiver sido dada e se pelo menos uma das características do item 1 for detectada, é necessário construir o gráfico novamente, mas agora sem as especificações, para verificar se o processo está realmente fora de controle, ou não consegue suprir as mesmas.
- 3- É importante ressaltar também que, quando detectada alguma característica do item 1 e o item 3 já tiver sido verificado, o próximo passo é eliminar as amostras que apresentaram esse item e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou existe um problema crônico e, assim, poder usar esses limites no futuro.

Quadro 4.1.a: Algoritmo de construção e análise do Gráfico *c* para amostras de mesmo tamanho.

A – Procedimentos iniciais

- 1- Definir qual a unidade amostral.
- 2- Estabelecer o número de amostras *m*.
- 3- Coletar as amostras na indústria ou estabelecimento em questão.
- 4- Calcular o total de defeitos, x_i , para cada amostra *i*.

B – Construindo os gráficos de acordo com suas características

B.1 – Parâmetro especificado

- 1- Construir o gráfico de x_i , no qual a abscissa são os indicadores de cada amostra, de 1 a *m*, e a ordenada são os valores do número de defeitos para cada

amostra.

2- Grafar a referência, cujo valor é dado por c .

3- Grafar os limites de controle $c \pm 3\sqrt{c}$.

B.2 – Parâmetro não especificado

1- Construir o gráfico de x_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores do número de defeitos para cada amostra.

2- Grafar a referência, cujo valor é dado por $\bar{c} = \frac{1}{m} \sum_{i=1}^m x_i$.

3- Grafar os limites de controle $\bar{c} \pm 3\sqrt{\bar{c}}$.

C – Análise dos gráficos

1- Verificar a existência dos três possíveis (principais) itens a seguir no gráfico construído, que são sinais de falta de controle do processo:

- Existem pontos fora dos limites de controle (isto é, acima do *LSC* ou abaixo do *LIC*)?
- Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo (sequências de pontos acima ou abaixo da linha central do gráfico)?
- Existem tendências ou sazonalidades?

2- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.

3- Se uma especificação tiver sido dada e pelo menos uma das características do item 1 for detectada, é necessário construir o gráfico novamente, mas agora sem a especificação, para verificar se o processo está realmente fora de controle, ou não consegue suprir a especificação.

4- É importante ressaltar também que, quando detectada alguma característica do item 1 e o item 3 já tiver sido verificado, o próximo passo é eliminar as amostras que apresentaram esse item e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou existe um problema crônico e, assim, poder usar esses limites no futuro.

Quadro 4.2.a: Algoritmo de construção e análise do Gráfico u para amostras de tamanhos iguais ou variáveis.

A – Procedimentos iniciais

- 1- Definir qual a unidade amostral.
- 2- Estabelecer o número de amostras m .
- 3- Coletar as amostras na indústria ou estabelecimento em questão.
- 4- Calcular o número médio de defeitos para cada amostra, dado por $\hat{u}_i = \frac{x_i}{n_i}$, em que x_i é o total de defeitos na i -ésima amostra e n_i é o tamanho da i -ésima amostra.

B – Construindo os gráficos de acordo com suas características

B.1 – Parâmetro especificado e não especificado (amostras de mesmo tamanho)

- 1- Construir o gráfico de \hat{u}_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores do número médio de defeitos para cada amostra.
- 2- Caso exista especificação, grafar a referência, cujo valor é dado por u . Senão, calcule $\bar{u} = \frac{1}{m} \sum_{i=1}^m \hat{u}_i$.
- 3- Grafar os limites de controle $u \pm 3\sqrt{\frac{u}{n}}$, se houver especificação, ou $\bar{u} \pm 3\sqrt{\frac{\bar{u}}{n}}$, se não houver.

B.2 – Parâmetro especificado e não especificado (amostras de tamanhos variáveis)

- 1- Construir o gráfico de \hat{u}_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores do número médio de defeitos para cada amostra.
- 2- Caso exista especificação, grafar a referência, cujo valor é dado por u . Senão, calcule $\bar{u} = \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m n_i}$.
- 3- Grafar os limites de controle $u \pm 3\sqrt{\frac{u}{n_i}}$, se houver especificação, ou $\bar{u} \pm 3\sqrt{\frac{\bar{u}}{n_i}}$,

se não houver.

4- Calcular para cada amostra o valor de z_i , considerando que $z_i = \frac{\hat{u}_i - u}{\sqrt{u/n_i}}$, caso u

tenha sido especificado, ou $z_i = \frac{\hat{u}_i - \bar{u}}{\sqrt{\bar{u}/n_i}}$, caso contrário.

5- Construir o gráfico de z_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores de z_i calculados para cada amostra.

6- Grafar o valor de referência, que deve ser igual a 0 e os limites de controle, que devem ser iguais a ± 3 .

C – Análise dos gráficos

1- Verificar a existência dos três possíveis (principais) itens a seguir no(s) gráfico(s) construído(s), que são sinais de falta de controle do processo:

- Existem pontos fora dos limites de controle, isto é, acima do *LSC* ou abaixo do *LIC*?
- Os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo, isto é, existem sequências de pontos acima ou abaixo da linha central do gráfico em um intervalo de tempo?
- Existem tendências ou sazonalidades no Gráfico u para amostras de mesmo tamanho ou no Gráfico u padronizado?

2- Se pelo menos um dos itens acima for verificado, deve-se investigar se existiram causas externas na coleta dos dados, a fim de evitar erros futuros na qualidade do processo.

3- Se uma especificação tiver sido dada e pelo menos uma característica do item 1 for detectada, é necessário construir o gráfico novamente, mas agora sem a especificação, para verificar se o processo está realmente fora de controle, ou não consegue suprir a especificação.

4- É importante ressaltar também que, quando detectada alguma característica do item 1 e o item 3 já tiver sido verificado, o próximo passo é eliminar as amostras que apresentaram esse item e reconstruir os gráficos de controle, para saber se apenas essas amostras foram responsáveis pela falta de controle ou existe um problema crônico e, assim, poder usar esses limites no futuro.

Quadro 6.1.a: Algoritmo de construção e análise dos gráficos de controle da variabilidade de múltiplas variáveis, para amostras com mais de uma observação.

1. Definir o tamanho e o número de amostras, representados por m e n , respectivamente.
2. Estabelecer o valor ideal para a matriz de covariâncias Σ , ou estimar esse valor através da matriz de covariâncias amostral S , calculada a partir de amostras coletadas em análises anteriores ou a partir das amostras disponíveis para a análise corrente.
3. Coletar os dados referentes às p variáveis que refletem a qualidade do processo.
4. Calcular, para cada amostra, a matriz dada por $A_i = (n-1)S_i$.
5. Calcular, para cada amostra, o valor da estatística dada por $W_i = -pn + pn \ln(n) - n \ln\left(\frac{|A_i|}{|\Sigma|}\right) + tr(\Sigma^{-1}A_i)$, sendo que tr representa o operador traço de uma matriz, $|\Sigma|$ representa o determinante da matriz Σ e \ln representa o logaritmo natural.
6. Construir o gráfico de W_i , no qual a abscissa são os indicadores de cada amostra, de 1 a m , e a ordenada são os valores de W_i calculados para cada amostra.
7. Representar no gráfico, através de uma reta horizontal, o valor do limite superior de controle, dado pelo α -ésimo quantil de uma distribuição Qui-quadrado com $\frac{p(p+1)}{2}$ graus de liberdade, ou seja, $\chi_{p(p+1)/2;\alpha}$.
8. Caso seja dada uma especificação para a matriz Σ e existam pontos acima do limite superior de controle do gráfico, deve-se construir o gráfico sem tal especificação de modo a determinar se houve alterações ao longo do tempo nos parâmetros do processo ou se o processo simplesmente não é capaz de atender às especificações estabelecidas.
9. Caso existam pontos que se encontram acima do limite superior de controle do gráfico sem especificações, deve-se investigar se houve alguma alteração nas condições de ocorrência do processo, que possa ter sido responsável pela

alteração ao longo do tempo no padrão de variabilidade das medidas observadas, ou se o gráfico apresenta apenas alarmes falsos para a falta de controle do processo.

10. Calcular o determinante da matriz de covariâncias de cada amostra, $|S_i|$.
11. Calcular o valor da estatística b_1 para cada amostra, sendo que, para a i -ésima amostra, esta estatística é dada por $b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i)$.
12. Calcular o valor da estatística b_2 para cada amostra, sendo b_2 , para a i -ésima amostra, dada por $b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^p (n-i) \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right]$.
13. Calcular o valor de referência do gráfico, dado por $b_1|\Sigma|$, caso Σ tenha sido especificado, ou $|S|$, caso contrário.
14. Calcular os limites de controle, dados por $|\Sigma|(b_1 \pm 3b_2^{0.5})$, caso Σ tenha sido especificado, ou $\frac{|S|}{b_1}(b_1 \pm 3b_2^{0.5})$, caso contrário.
15. Construir o gráfico da série temporal calculada no item 10.
16. Representar no gráfico, através de uma reta horizontal, o valor de referência calculado no item 13.
17. Representar no gráfico, através de duas retas horizontais, os limites de controle calculados no item 14.
18. Verificar se a série apresenta qualquer uma das seguintes características: (i) há pontos que se encontram fora dos limites de controle; (ii) os pontos não se distribuem aleatoriamente em torno da linha central em um intervalo de tempo; (iii) há tendências ou padrões cíclicos (sazonalidades).
19. Caso qualquer uma das três características descritas no passo anterior seja observada, deve-se investigar se há causas que possam ser consideradas como responsáveis por tais problemas, a fim de evitar erros futuros na qualidade do processo. Para facilitar tal investigação, recomenda-se construir o Gráfico S ou o Gráfico R , de cada variável, para identificar melhor quais as variáveis que são as maiores responsáveis pela potencial falta de controle do processo.

20. Se uma especificação tiver sido dada e qualquer dos problemas do item 18 tiver sido observado, deve-se construir o gráfico de controle sem considerar a especificação dada e verificar se o problema persiste, para saber se o processo está de fato fora de controle ou se simplesmente não é capaz de atender à especificação estabelecida.
21. Se nenhuma especificação para o processo tiver sido dada, houver interesse em usar em análises futuras os limites de controle calculados através das amostras disponíveis atualmente, qualquer dos problemas listados no item 18 tiver sido observado e se forem identificadas as causas responsáveis por tais problemas, o próximo passo é excluir todas as amostras que sofreram efeitos de tais causas e calcular novamente os limites de controle.

Quadro 6.2.a: Algoritmo de construção e análise do Gráfico T^2 para amostras com mais de uma observação (dados agrupados).

1. Determinar o tamanho e o número de amostras, representados por n e m , respectivamente.
2. Coletar os dados referentes às p variáveis que refletem a qualidade do processo, sendo $\mathbf{x}'_{ij} = (x_{1ij}, x_{2ij}, \dots, x_{pij})$ definido como o vetor de variáveis referente à j -ésima observação da i -ésima amostra, para $i=1, 2, \dots, m$ e $j=1, 2, \dots, n$.
3. Calcular o vetor de médias de cada amostra,

$$\bar{\mathbf{x}}'_i = \left(\frac{1}{n} \sum_{j=1}^n x_{1ij}, \frac{1}{n} \sum_{j=1}^n x_{2ij}, \dots, \frac{1}{n} \sum_{j=1}^n x_{pij} \right).$$
4. Prosseguir com os passos 5 a 11, se tiverem sido dadas especificações para o processo, ou com os passos 12 a 18, caso contrário.
5. Estabelecer os valores ideais ou os valores que se supõe que sejam os verdadeiros para o vetor de médias $\boldsymbol{\mu}$ e para a matriz de covariâncias Σ das variáveis observadas.
6. Calcular, para cada amostra, a estatística dada por $T_i^2 = n(\bar{\mathbf{x}}'_i - \boldsymbol{\mu})' \Sigma^{-1} (\bar{\mathbf{x}}'_i - \boldsymbol{\mu})$.
7. Calcular o limite superior de controle, dado pelo α -ésimo quantil de uma distribuição Qui-quadrado com p graus de liberdade, $\chi^2_{p;\alpha}$. Estabelecer o limite inferior de controle como sendo igual a zero.

8. Construir o gráfico da série temporal dos T_i^2 .
9. Representar no gráfico, através de duas retas horizontais, os limites de controle calculados no item 7.
10. Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se verificar se tais amostras sofreram efeito de alguma alteração nas condições de ocorrência do processo, visando saber se o processo se encontra realmente fora de controle ou se ocorreram, no caso de tais amostras, apenas alarmes falsos.
11. Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se ainda construir o Gráfico T^2 sem especificações, utilizando os passos que se seguem, de forma a determinar se houve, por um período de tempo, uma alteração no vetor de médias das variáveis observadas ou se o verdadeiro vetor de médias do processo apenas difere do estipulado *a priori*.
12. Estimar o vetor de médias do processo através das amostras disponíveis para a análise corrente ou através de um número grande de amostras utilizadas em análises anteriores. A estimativa desse vetor é dada por
- $$\bar{\mathbf{x}}' = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p) = \left(\frac{1}{m} \sum_{i=1}^m \bar{x}_{1i}, \frac{1}{m} \sum_{i=1}^m \bar{x}_{2i}, \dots, \frac{1}{m} \sum_{i=1}^m \bar{x}_{pi} \right).$$
13. Estimar a matriz de covariâncias do processo através das mesmas amostras usadas no passo anterior. Essa estimativa é dada por
- $$S = \begin{bmatrix} \bar{S}_1^2 & \bar{S}_{12} & \dots & \bar{S}_{1p} \\ \bar{S}_{12} & \bar{S}_2^2 & \dots & \bar{S}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{S}_{1p} & \bar{S}_{2p} & \dots & \bar{S}_p^2 \end{bmatrix}, \text{ sendo que } \bar{S}_k^2 \text{ representa a média das variâncias}$$
- amostrais referentes à k -ésima variável e \bar{S}_{kg} representa a média das covariâncias amostrais entre a k -ésima e a g -ésima variável, $k \neq g$.
14. Calcular, para cada amostra, a estatística dada por $T_i^2 = n(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' S^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})$.
15. Calcular o limite superior de controle. Se o interesse for determinar quais são as

amostras fora de controle, este limite é dado por $\frac{p(m-1)(n-1)}{mn-m-p+1} F_{p, mn-m-p+1; \alpha}$.

Se as amostras fora de controle já foram retiradas do banco de dados e pretende-se utilizar as demais amostras para se determinar os valores de referência para os gráficos a serem utilizados em análises posteriores (as estimativas $\bar{\bar{x}}$ e S e os limites de controle) visando usar tais estimativas em

análises posteriores, este limite é dado por $\frac{p(m+1)(n-1)}{mn-m-p+1} F_{p, mn-m-p+1; \alpha}$.

Assumir o limite inferior de controle igual a zero.

16. Construir o gráfico da série temporal dada pela estatística T_i^2 , calculada no passo 14.
17. Representar no gráfico, através de duas retas horizontais, os limites de controle calculados no passo 15.
18. Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se verificar se tais amostras sofreram efeito de alguma alteração nas condições de ocorrência do processo, visando saber se o processo se encontra realmente fora de controle ou se ocorreram, no caso de tais amostras, apenas alarmes falsos.

Quadro 6.2.b: Algoritmo de construção e análise do Gráfico T^2 para amostras com apenas uma observação (dados individuais).

1. Determinar o número de amostras, m , a serem coletadas.
2. Coletar os dados referentes às p variáveis que refletem a qualidade do processo, sendo $\mathbf{x}'_i = (x_{1i}, x_{2i}, \dots, x_{pi})$ definido como o vetor de variáveis referente à i -ésima amostra, para $i=1, 2, \dots, m$.
3. Prosseguir com os passos 4 a 10, se tiverem sido dadas especificações para o processo, ou com os passos 11 a 17, caso contrário.
4. Estabelecer os valores ideais ou os valores que se supõe que sejam os verdadeiros para o vetor de médias $\boldsymbol{\mu}$ e para a matriz de covariâncias Σ das variáveis observadas.
5. Calcular, para cada amostra, a estatística dada por $T_i^2 = (\mathbf{x}_i - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$.

6. Calcular o limite superior de controle, dado pelo α -ésimo quantil de uma distribuição Qui-quadrado com p graus de liberdade, $\chi_{p,\alpha}^2$. Estabelecer o limite inferior de controle como sendo igual a zero.
7. Construir o gráfico da série temporal dos T_i^2 .
8. Representar no gráfico, através de duas retas horizontais, os limites de controle calculados no item 6.
9. Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se verificar se tais amostras sofreram efeito de alguma alteração nas condições de ocorrência do processo, visando saber se o processo se encontra realmente fora de controle ou se ocorreram, no caso de tais amostras, apenas alarmes falsos.
10. Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se ainda construir o Gráfico T^2 sem especificações, utilizando os passos que se seguem, de forma a determinar se houve, por um período de tempo, uma alteração no vetor de médias das variáveis observadas ou se o verdadeiro vetor de médias do processo apenas difere do estipulado *a priori*.
11. Estimar o vetor de médias do processo através das amostras disponíveis para a análise corrente ou através de um número grande de amostras utilizadas em análises anteriores. A estimativa desse vetor é dada por

$$\bar{\mathbf{x}}' = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p) = \left(\frac{1}{m} \sum_{i=1}^m x_{1i}, \frac{1}{m} \sum_{i=1}^m x_{2i}, \dots, \frac{1}{m} \sum_{i=1}^m x_{pi} \right).$$
12. Estimar a matriz de covariâncias através de $S = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$ ou através de $S = \frac{1}{2(m-1)} V'V$, em que $V = \begin{bmatrix} (\mathbf{x}_2 - \mathbf{x}_1)' \\ (\mathbf{x}_3 - \mathbf{x}_2)' \\ \dots \\ (\mathbf{x}_n - \mathbf{x}_{n-1})' \end{bmatrix}$.
13. Calcular, para cada amostra, a estatística dada por $T_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})' S^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$.

- 14.** Calcular o limite superior de controle. Se o interesse for determinar quais são as amostras fora de controle, este limite é dado por $LSC = \frac{(m-1)^2}{m} \beta_{p/2, (m-p-1)/2; \alpha}$.
- Se as amostras fora de controle já foram retiradas do banco de dados e pretende-se utilizar as demais amostras para se determinar os valores de referência para os gráficos a serem utilizados em análises posteriores (as estimativas \bar{x} e S e os limites de controle), visando usar tais estimativas em análises posteriores, este limite é dado por $LSC = \frac{p(m+1)(m-1)}{m^2 - mp} F_{p, m-p; \alpha}$.
- Assumir o limite inferior de controle como sendo igual a zero.
- 15.** Construir o gráfico da série temporal dada pela estatística T_i^2 , calculada no passo 13.
- 16.** Representar no gráfico, através de duas retas horizontais, os limites de controle calculados no passo 14.
- 17.** Caso existam amostras que apresentem uma estatística T_i^2 maior do que o limite superior de controle, deve-se verificar se tais amostras sofreram efeito de alguma alteração nas condições de ocorrência do processo, visando saber se o processo se encontra realmente fora de controle ou se ocorreram, no caso de tais amostras, apenas alarmes falsos.